HEMICAL REVIEWS

W. A. NOYES, EDITOR-IN-CHIEF

*Professor of Chemistry and Director of the Chemical Laboratory
of the University of Illinois*

VOLUME IV

BALTIMORE

1927

CONTENTS

COLLOID DYNAMICS[1]

VICTOR COFMAN

Wilmington, Delaware

"Simplicity and symmetry should be among the chief aims of a scientific theory. It is probable that the same laws which regulate the movement of electrons within the atom also determine the paths of planets in their orbits; a complete understanding of the simplest phenomenon may enable us to explain the Universe."

ABSTRACT

Several phenomena, at present little known or neglected, are described and explained: spontaneous dispersion of liquids, change in surface tension of solutions with pH; transformation of chemical into surface energy, "thermo-osmosis," etc.

New experimental methods are given for the determination of (a) swelling pressure of gels, (b) concentration of soap in films, (c) velocity of adsorption at liquid surfaces.

The similarity between the following variables and phenomena is discussed:

P' = Colloid pressure;   Osmotic pressure
    Swelling pressure of gels;
    Surface pressure of adsorbed
    molecules.

V' = Volume of gel;
    Volume of Surface Region. V = Volume of solution

T' = Colloid potential; T = Temperature, or heat potential.
    pH, in special cases.

E' = Colloid energy; E = Heat energy or molecular energy.
    Electro-chemical energy.

S' = Colloid entropy: $\int \dfrac{dE'}{T'}$ S = Heat entropy: $\int \dfrac{dE}{T}$

1

M' = Mass of oriented molecules

F' = Partial molal free energy, = $\dfrac{dE'}{dM'}$.

Adsorption at surfaces

Spontaneous dispersion
Lyophobic and lyophyllic colloids
Electro-osmosis
Stream Potential
Electro-phoresis

Colloid Engine;
Muscular System
Propagation of nerve stimuli

M = Mass of solute molecules.

F = Partial molal free energy $\dfrac{dE}{dM}$

Distribution of crystalloid between two phases.
Boiling of liquids
Vapors and permanent gases
"Thermo-osmosis"
"Stream Temperature"
"Thermo-phoresis";
Crook's Radiometer
Heat Engine

Propagation of disturbances in elastic media, e.g. sound waves.

The parallelism in most cases is complete, the relations between the colloid variables P', V', T', etc., being identical with those holding between the "crystalloid" or "gas" variables, P, V, T, etc. The fundamental relation for crystalloids or colloids is

$$dE = TdS - PdV + F_n dM_n$$
$$= T'dS' - P'dV' + F_n dM_n$$

or, in general, introducing t, the time variable, and equating to zero:

$$F_n dM_n - P_n dV_n + T_n dS_n - X_n dt_n = 0$$

This is a symmetrical equation in which X stands for dE/dt, or "power" function, and n for any number of variables of the same type: M_n may represent the mass of gas, colloid, solute, electrons, etc.; P_n may stand for gas pressure, colloid pressure, osmotic pressure, electromotive force, and so on. The relation between entropy and time is similar to that between mass and volume; in other words, *entropy extends in time* just as *mass extends in space*. Or, the events of which our world is composed may be separated into a space-component, mass, and a time-component, entropy.

This point of view leads to a simple thermodynamic distinction between animate and inanimate systems: in a non-living process entropy may be considered at rest or in simple motion; in a liv-

ing system entropy is in complex motion; by comparing the two, a quantitative measure of the life factor may be obtained. Systems containing catalysts or enzymes are apparently of an intermediate type in which the displacement of entropy in time is of a predictable character.

* * * *

WHAT IS A COLLOID?

The original definition of colloids endows them with negative characteristics only; they *are not* crystalline, *do not* diffuse through membranes, *do not* exert osmotic pressure and *do not* affect the vapor pressure or the freezing point of water. Apart from being of an entirely negative character, these statements are true only in a very restricted sense; many crystalline substances can exist in the colloidal state, the vapor pressure and freezing point of gels differ greatly from those of their liquid, and the swelling pressure is not distinguishable in its action from osmotic pressure.

The size of the particles in a system has been used by Zsigmondy (1) and others as a criterion of colloidality. According, to this view, colloid systems contain particles much larger than molecules, yet below the limit of microscopic vision (10^{-5} to 10^{-7} cm. diam.). Wo. Ostwald has stressed the necessity of several phases being present in a colloid system and has pointed out the large amount of surface accompanying the state of fine dispersion. At the present time the orientation theory emphasizes the fact that colloid properties depend on a certain arrangement of the molecules; at the boundary of a phase, orientation always takes place to a greater or less extent. In elastic gels the orientation may be considered to occur in the interior of a liquid, independent of the presence of an interface, unless the oriented molecules themselves be considered to represent another phase (Bancroft (2)).

In this paper we shall follow the thermodynamic method, which is independent of any structural theory. Nevertheless, it will make the argument clearer if we indicate occasionally the picture we have in mind when discussing a given phenomenon. *It is convenient to imagine* that a system will have colloidal prop-

erties if it contains large polar molecules, or particles, more or less oriented. Such particles possess little or no energy of motion. There is, then, at one extreme the perfect gas composed of molecules which move haphazardly; its energy is a function of the temperature only. At the other extreme is the ideal colloid, a substance with large oriented molecules which have but little freedom of movement. The energy of the colloid is not, therefore, heat energy (irregular molecular motion), but another form of energy which will be shown later to be of an electrochemical nature.

The various definitions of colloids merely look upon the subject from different points of view: A high degree of dispersion (Zsigmondy, Alexander, Von Weimarn (3)) implies a large amount of interface (Ostwald), which in turn leads to the orientation of molecules (Harkins, Langmuir, Hardy (3)) and involves the presence of energy other than heat (Einstein (4)).

Colloids, it may be added, are not chemical compounds in the strict sense of the word. Their composition varies continuously with changes in physical conditions (see section on colloid energy). It would be convenient to designate them by the name *physical compounds*.

THE COLLOID VARIABLES

The state of a gas or of an ordinary solution depends upon various factors, such as pressure, volume, and temperature. By the application of certain fundamental principles which determine the relation between those factors, a great advance in our physico-chemical knowledge has taken place during the last few decades. If, however, we try to apply to colloids the thermodynamic relations which have been found so useful in the case of ordinary solutions, we are confronted with a difficult situation, because P, V, T, are no longer important variables where colloids are concerned. P, the osmotic pressure is, by definition, zero or negligible in the case of colloid solutions; in place of volume it is the surface which is important. The temperature too, must often be kept constant in a colloid system. This is shown by the fact that the colloids which form our body, and

which probably come nearest to the ideal state, must be maintained at a constant temperature. All this is in accordance with our picture of the colloid as a substance in which energy is present in a form other than heat energy.

Since the thermodynamic variables, P, V, T, etc., are not suitable, we shall find another set of variables to take their place. For simplicity we shall denote these new variables by the corresponding letters, P′, V′, T′, etc. Like ordinary pressure, volume, and temperature, these new variables are magnitudes that can be measured by experiment, and are *independent of any theories*.

The colloid pressure P′: The swelling pressure of gels. When a sugar solution is enclosed in a semipermeable membrane, such as parchment paper, and immersed in pure water, the water penetrates into the sugar solution and builds up a certain pressure. This is called osmotic pressure. If a gelatine gel be used in place of sugar solution, it behaves in the same way; it absorbs water and exerts pressure, the so-called pressure of swelling of gels.

The great pressure exerted by colloids on swelling has been known from ancient times. The Egyptians used the pressure of swelling of wood to dislocate huge blocks of stone. Conversely, when shrinking owing to loss of water, gelatine pulls with sufficient force to chip the glass to which it is attached.[2]

Posnjack (5) has measured directly, in an osmotic cell, the pressure of swelling of gelatine, up to 6 atmospheres (fig. 1). The swelling pressure of gelatin may also be measured by balancing it against the osmotic pressure of crystalloid solutions. If a series of cubes of a gel (containing, say, 25 per cent dry gelatine) are placed in sugar solutions of different strengths, then the gelatine acts to a large extent as its own semipermeable membrane and according to the concentration of the surrounding sugar, water passes from the gel to the solution or vice versa (fig. 2).

---

[2] The colloid pressure, it may be remarked, offers a simple mechanism for the rise of sap in plants. It is known that osmotic pressure due to the crystalloid substances in the plant cells is insufficient to account alone for the pressure necessary for that purpose.

Posnjak's original data were expressed in terms of concentration of gelatine per 100 grams of gel (fig. 1) and an exponential
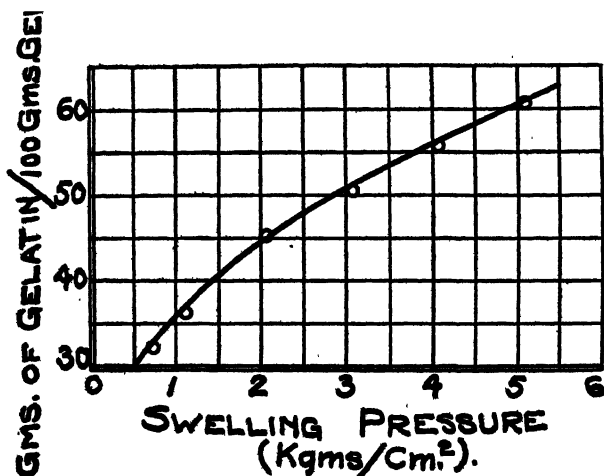
FIG. 1. POSNJAK'S DATA ON THE SWELLING PRESSURE OF GELATINE
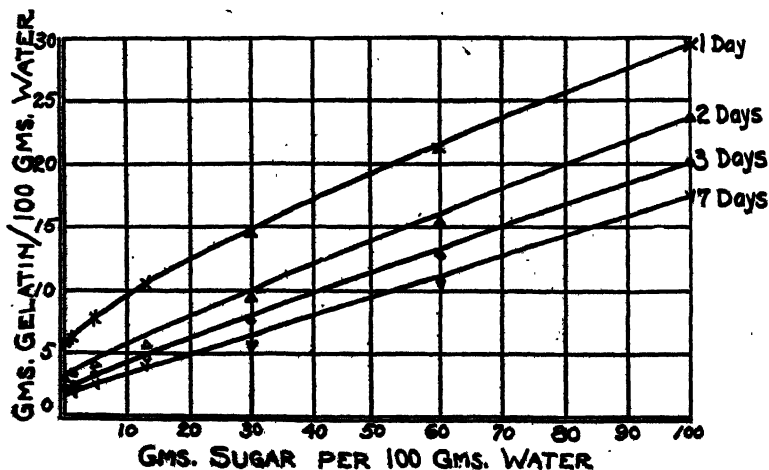
FIG. 2. EQUILIBRIUM BETWEEN OSMOTIC PRESSURE AND SWELLING PRESSURE

equation was fitted to represent the relation between concentration and swelling pressure, namely $P = P_0 C^n$ (n = nearly 3).

It is customary, on the other hand, to express osmotic pressure results in relation to concentration of solute per 100 grams of solvent (6). By plotting Posnjak's data in the same way (fig. 3) it will be seen that the curves become much straighter than in figure 1. Further, it was assumed in deducing the exponential formula that C = 0 when P = 0; but it is well known that [(below 20°) a gel will come to equilibrium with pure, or nearly pure solvent. This means that a formula of the type P = AC + const. would be more consistent with facts.
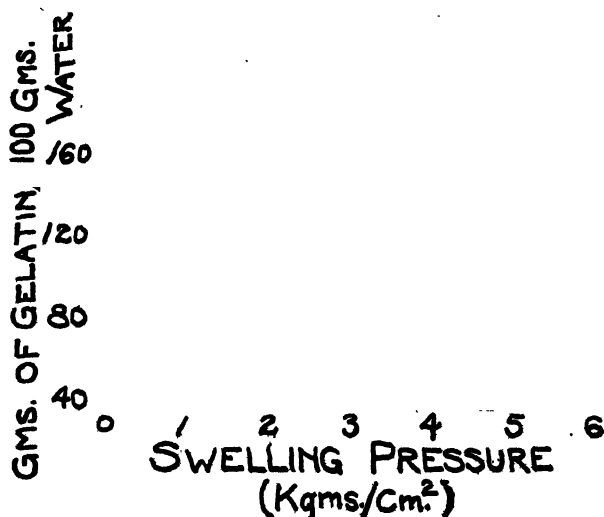
FIG. 3. DATA OF FIGURE 1 RECALCULATED IN TERMS OF CONCENTRATION OF GELATIN PER 100 GRAMS OF WATER

The experiments with sugar solution strongly suggest that whatever law connects concentration and osmotic pressure, also connects concentration and swelling pressure of gels. These experiments will have to be repeated with a real semipermeable membrane between the sugar and gelatin, in order to prevent completely the diffusion of sugar, before final conclusions can be drawn.

The general mathematical relations which we shall presently develop are independent of the particular law which governs

pressure and concentration. Even when the "gas laws" are used, a linear relation need not be presupposed, if the "activity coefficient" (α) is introduced.

The surface pressure of adsorbed films. When a film of certain substances is present at the surface of a liquid, it exerts a pressure. This pressure is measured experimentally as the difference between the surface tension of the pure liquid and the surface tension of the surface contaminated with the adsorbed substance. Figure 4 shows a simple apparatus which may be

SURFACE PRESSURE APPARATUS.
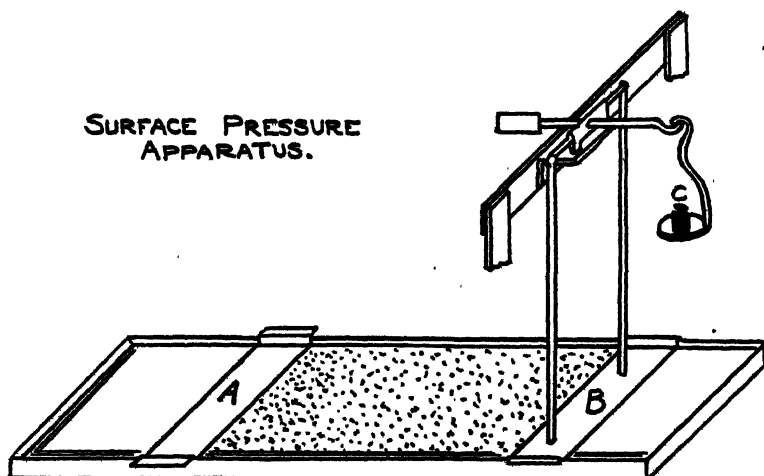
FIG. 4. SURFACE PRESSURE APPARATUS

used for the purpose: a shallow tray filled with water having a fixed barrier (paraffined copper plate or rod) at A, and a movable aluminum "float" at B. A monomolecular film of, say, oleic acid, lying on the water between A and B exerts a pressure on the float B, and this can be measured by placing weights on the balance at C. A more complete account of the behavior of thin films will be found in an article by N. K. Adam in this journal (7).

The colloid volume V' is the space within which the pressure P' is active. In gels V' may be taken to be the volume of the gel, or more correctly, the volume of the gel less the volume actually

occupied by the colloid. This corresponds to the (v − b) factor in van der Waals' equation of state for gases, and is necessary in order to obtain a simpler relation between pressure and volume, both in ordinary solutions and in gels (6).

In surface solutions V' is the volume of the surface region. In figure 5, XY represents the surface of a liquid which, when pure, has a surface tension σ represented by the inward pointed arrows. An adsorbed substance at the surface exerts an outward pressure, F, which opposes the surface tension. This pressure is similar to the osmotic pressure of ordinary solutions. It has been customary to express this "pressure" F in *dynes per cm.* The pressure, however, acts over a small but finite distance,
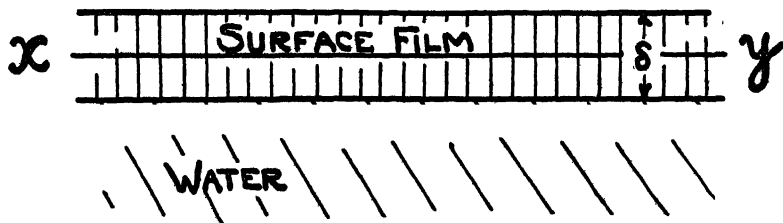
FIG. 5. OPPOSITE DIRECTIONS OF VECTORS FOR SURFACE PRESSURE AND SURFACE TENSION

δ, which is the thickness of the surface region. F/δ therefore represents the surface pressure and Aδ the volume of the surface region, A being the area.

Returning for a moment to our structural idea, we may imagine that the concentration of oriented molecules becomes so great that they can "join hands," so to say, and form a continuous frame-work (8). Under those circumstances, when all the solvent is within the region where the force of orientation acts, the system becomes an elastic gel. On the assumption that the surface pressure and the swelling pressure are identical, we may write P' = F/δ. On the same assumption, we may calculate the value of the thickness of the surface region, from the concentration of colloid per cm. square of surface, Γ, and from the minimum concentration of colloid that will give a gel C_m, because, assuming

complete molecular orientation, we have $\delta = \dfrac{\Gamma}{C_m}$. The transition from gel to sol, however, is not sharp and δ will not be an exactly defined quantity; the result, nevertheless, may serve to indicate the order of magnitude of the surface region.

In the case of sodium oleate, C_m, the minimum concentration needed to give a gel is 0.24 grams per cubic centimeter (9). The "area per molecule" in a mono-molecular soap film has been calculated by Harkins and Zollman (10) from emulsion experiments to be about 47×10^{-16} cm.^{2}, which corresponds roughly to 1×10^{-7} grams of sodium oleate per square centimeter of interface. The writer, by a direct method, has found for the concentration of soap per square centimeter of foam surface a value of approximately 2×10^{-7} grams. The thickness of the surface layer is therefore of the order of magnitude

$$\frac{1.5 \times 10^{-7}}{0.24} = 6 \times 10^{-7} \text{ cm.}$$

or 60 Ångstrom units. The length of the oleic molecule itself has been computed to be 11 to 27 Å (from the thickness of surface films (11) and from x-ray measurements (12)).

The direct method for determining the concentration of sodium oleate in foam is as follows:

One cubic centimeter of a solution containing about 0.1 per cent sodium oleate and an equal amount of sodium carbonate (to keep the pH constant at about 10) is placed in a small glass tube and air free from CO_2 is forced through a capillary so as to form uniform bubbles of about 0.05 cm. diameter. Under these conditions, practically all the soap can be obtained in the form of a permanent foam of uniform bubbles. The surface area of this foam can be readily calculated, being equal to

$$\frac{\text{volume} \times 6}{\text{diameter of bubbles}}$$

In a series of experiments at 22°, 1 cc. of N/500 solution containing 0.06 per cent sodium oleate yielded on the average 35 cc.

of foam with uniform bubbles of 0.07 cm. diameter, representing a surface of

$$6 \times 35 \div 0.07 = 3000 \text{ cm.}^2$$

therefore the *maximum* amount of sodium oleate per square centimeter of surface is

$$\frac{0.0006}{3000} \quad 2 \times 10^{-7} \text{ gms.}$$
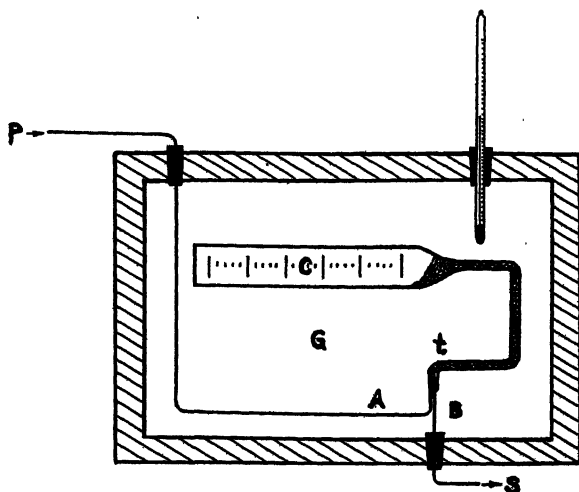
FIG. 6. APPARATUS FOR DETERMINING THE CONCENTRATION OF SOAP IN FOAM

Corrections would have to be made for the amount of soap left in solution and for that present in the interior of the film, but these would both tend to decrease the value of δ. The above method can be employed for determining the foaming power of soaps in absolute units. Figure 6 shows how the apparatus can be arranged for convenient use:

The tube t containing the soap solution is connected with two capillaries, A and B. Through A air is introduced under a definite pressure, measured by the manometer attached to the inlet tube at P. The capillary B serves to introduce the solution at the beginning, and to withdraw it at the end of the experiment,

also to clean the apparatus without having to remove it from the thermostat, *G*.

The writer wishes to emphasize the fact that the experiments described here, and elsewhere in this paper, were carried out in an industrial laboratory and that no attempt has been made to secure a high degree of accuracy. They are given here for the sake of the principles involved and not as exact quantitative determinations. The results obtained by other investigators are given for comparison wherever data are available.
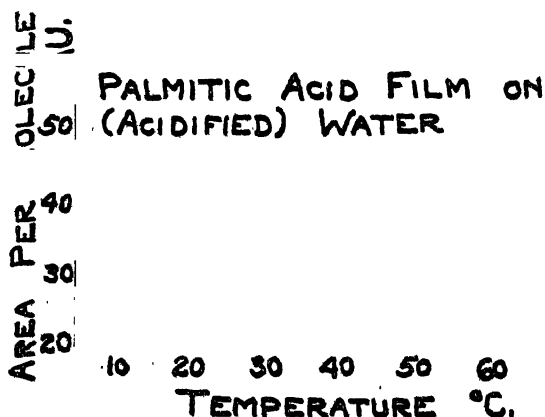
FIG. 7. INFLUENCE OF TEMPERATURE ON SURFACE FILM

(N. K. Adam. Proc. Roy. Soc. 101, (A), 519 (1922))

The colloid "temperature" or potential T'. It has been stated above that the temperature should be considered a constant in colloid systems. By this it is not meant that temperature does not affect colloids, but only that the changes caused by heat are discontinuous. Figure 7 shows the effect of temperature on the area occupied by a surface film of palmitic acid. (13). When the temperature rises, little or nothing is observed up to a certain point, in the neighborhood of 25° to 30°, when a sudden change takes place; after that, further increase in temperature has but slight effect. Evidently in this case the relation between temperature and area (or colloid volume) cannot be expressed by a simple formula.

Consider now figure 8 which shows how the OH⁻ion concentration (pH = log OH⁻ion concentration + 14) affects the surface pressure of fatty acid films (14). (The latter is roughly proportional—or symbat—to the drop number). It is evident that there is a gradual, and in some cases almost a linear change. The surface pressure can therefore be expressed as a simple function of the pH.
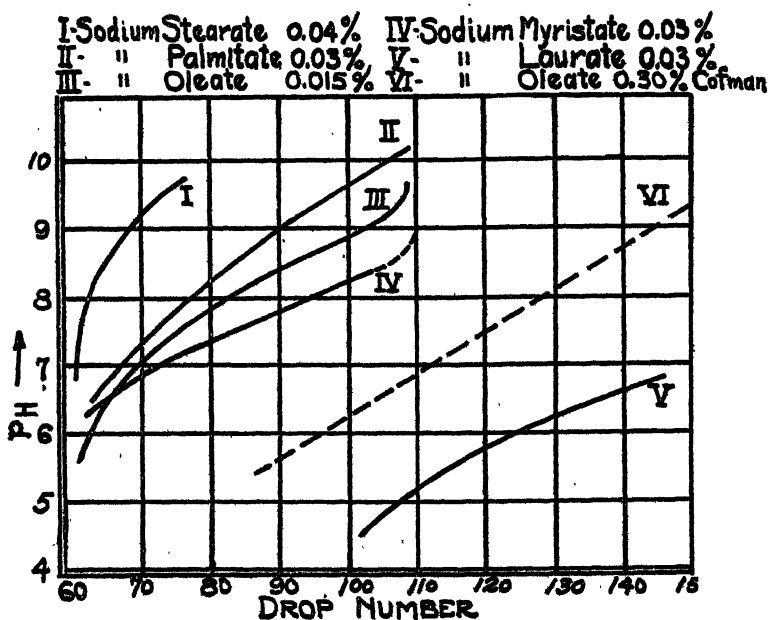
FIG. 8. INFLUENCE OF pH ON SURFACE TENSION OF SOAP SOLUTIONS

(S. Jarisch. Biochem. Zt. 134, 163 (1922))

The variable T which causes the pressure and volume of gases to increase is called temperature or heat potential. The variable T' which causes an increase in the colloid (surface) pressure and volume we shall denote by the name colloid potential. In the particular instance of fatty acid films, the pH is, within limits, a measure of the colloid potential, but it should be understood that this is not necessarily true for films of other substances.

Temperature is measured in many ways: by the expansion

of mercury, or of gas, by means of thermo-couples, radiation pyrometers, etc. Similarly, according to circumstances, in order to measure colloid potential we shall use the expansion of films and gels, hydrogen electrodes, color indicators, etc. If we had an ideal colloid whose volume V' diminished continuously with decreasing colloid potential, finally becoming zero, then the colloid potential corresponding to zero volume would be the absolute zero of the colloid potential. Like the perfect gas (15) "the ideal colloid is an invented substance, defined by certain properties which are not possessed by any actual substance, but which are supposed to be approached by many actual colloids (at great concentration and high potential). We may state, then, that the perfect colloid is a substance which fulfills the two following conditions:

1. That its energy is a function of the colloid potential alone, or in other words that $\left(\dfrac{\delta\ E'}{\delta\ V'}\right)_{T'} = 0$.

2. That when its T', P', V', are changed, these obey the relation $P'\ V' = M'\ R'\ T'$ where M' is the mass of colloid."

Following the example of thermodynamics, we shall define later the colloid potential as a work function and also show its connection to entropy.

Spontaneous dispersion. An understanding of the factors described in the previous paragraphs, namely, colloid pressure, volume and potential, enables us to explain the phenomenon of spontaneous dispersion of liquids. At the plane of contact of two immiscible liquids a few minute particles of one of the liquids (diameter of the order 1×10^{-5} cm.) are generally seen floating in the other phase. The dispersion is greatly increased by the presence of certain electrolytes and in some cases it becomes so intense that an emulsion is produced. This happens, for instance, when a 5 per cent solution of ferric chloride in nitrobenzol comes into contact with water; the water enters the nitrobenzol in the form of minute droplets which, under a magnification of 100 diameters, are seen to be in violent movement. If a membrane of parchment paper be used to separate the two liquids water still penetrates into the nitrobenzol building up a certain

pressure, as in the case of osmotic experiments. The emulsion thus formed is more stable than that obtained when no membrane is used.

Spontaneous emulsification of drops of rancid olive oil in alkaline water solution has been described by Gad (16). Maday (17) has observed the dispersion of oleic acid floating on aqueous ammonia. Gad also investigated the best conditions for the emulsification of oils containing fatty acids in aqueous solutions of sodium carbonate. The dispersion in these instances has been rightly attributed to the interaction between the fatty acid and the alkali, resulting in the formation of soap, which disturbs the equilibrium at the interface. According to Freundlich (18): "The soap formed at the interface by the interaction of fatty acid and alkali strongly depresses the interfacial tension and causes the formation of small drops. These drops do not coalesce, since the layer of soap at the interface acts as protective colloid. Indeed, the adsorbed layer of soap favors every increase in the interface; it fixes it, as it were; for the interfacial tension is small, and when by mechanical means depressions and contractions of form are produced in the liquid to be dispersed, they do not, on account of the low tension, disappear again. Indeed, one may, with Donnan, assume an influence which actually opposes the coalescence of the drops, and even divides them up further."

From the point of view of our new colloid variables, we are dealing here with a phenomenon analogous to the boiling of liquids. Oleic acid is adsorbed at the water-oil interface and exerts a pressure, lowering considerably the interfacial tension. (10). When it comes into contact with NaOH its potential increases, its pressure becomes greater than the interfacial tension, and dispersion occurs, just as a liquid boils when its vapor pressure becomes greater than the external pressure.

The energy needed for increasing the pressure of the adsorbed molecule is supplied in this case by the reaction between the acid and base. A chemical reaction does not seem to be always necessary, because in some cases good dispersion is obtained with one solute only: e.g., sodium benzoate in nitrobenzol and water. The energy is then obtained simply by the passage of a

substance from one phase into another in which its potential energy is smaller.

It can be readily shown by a rough calculation that the energy from the reaction between the fatty acid adsorbed at the interface and the alkali is sufficient to account for the surface energy formed during emulsification.

The amount of oleic acid adsorbed per square centimeter of surface (in the form of a monomolecular film) is about 1.5×10^{-7} gm. When this combines with alkali to form soap it liberates about 300 ergs of energy, which is certainly more than the surface energy per square centimeter of interface in an emulsion (the total surface energy of water is 118 ergs / cm.^{2} and that of nitrobenzol 77 ergs/cm.^{2} (19)).

The subject is amenable to exact treatment as follows: Let the oleic acid be dissolved in nitrobenzol, and let

E_1 = energy of reaction (H Ol) Nb + (Na OH) Aq.

H = heat liberated during spontaneous emulsification; then

λ_s = energy used in surface formation = $E_1 - H$.

E_1 can be calculated from

H_1 = heat of reaction (H Ol)Aq. + (Na OH) Aq.

H_2 = heat of reaction H Ol + Nb.

H_3 = heat of reaction H Ol + Aq.

because $E_1 = H_1 + H_2 - H_3$.

λ_s can also be calculated from the formula $\lambda_s = T^1 \dfrac{dP'}{dT'}$ which corresponds to Clapeyron's equation for the latent heat of evaporation. $\dfrac{dP'}{dT^1}$ is the change in surface pressure with pH, provided the latter may be taken as a measure of the colloid potential T' at the interface. The practical difficulties would probably consist in determining the extent of surface formed during emulsification.

It should be noted that λ_s, the energy of emulsification in the presence of a colloid, is *not* the same as the energy of formation of a new surface, which is given by the Kelvin equation,

$$\left(\frac{\delta E}{\delta A}\right)_T = \sigma - T\left(\frac{\delta\sigma}{\delta T}\right)_A$$

The former corresponds to the total heat of vaporization $T \frac{dP}{dT}$ (vapor), while the latter represents the energy used in the formation of a vacuum

$$\left(\frac{\delta E}{\delta V}\right)_T = P - T \left(\frac{\delta P_{gas}}{\delta T}\right)_V$$

δV being the volume of the vacuum formed and P the external pressure.

The colloid engine. The colloid potential T' has already been defined in terms of the expansion of the ideal colloid, just as temperature is often defined as a function of the expansion of the perfect gas. There is, however, a more satisfactory definition of temperature, as a work function: *It is possible to obtain work whenever a difference in temperature exists.* This is the principle on which all heat engines are designed. It can be shown that the colloid potential T' may be used as a work function in exactly the same manner. The apparatus for measuring surface tension, shown in fig. 4. will serve to show the principle of a colloid engine which transforms chemical energy into mechanical work, at constant temperature, given a difference in colloid potential.

If the film of oleic acid on the surface of the water, between A and B, has a concentration of one molecule per 40×10^{-16} cm.², it exerts a pressure of about 17 dynes/cm. on the float B (11). On adding carefully a drop of alkali to the film, or bringing over it some ammonia gas, the oleic acid is changed into soap and the pH (or — log. H-ion concentration) of the surface rises from 1.3 to 9.5, approximately. At the same time the pressure P' increases from 17 to 44 dynes/cms. which is the surface tension lowering due to a saturated film of soap. This, of course, is only a momentary change, because the soap formed soon reacts with the acid in the interior of the liquid and the film returns to its original state.

The system shown in Figure 4 can therefore be made to work as an engine, the strip of aluminum B, being the movable piston. With every increase in pH (addition of alkali) the pressure P' of the film increases and the piston is moved from B to a new

position C.   On removing the source of alkalinity the pH falls and the piston resumes its former position.

This represents a "colloid" engine in its simplest form.   In order to deduce the efficiency of the above "colloid" engine, we may take it through a reversible cycle, by making the external pressure on the piston B differ from P' only by an infinitely small amount dP, and assuming the usual frictionless piston and complete insulation.
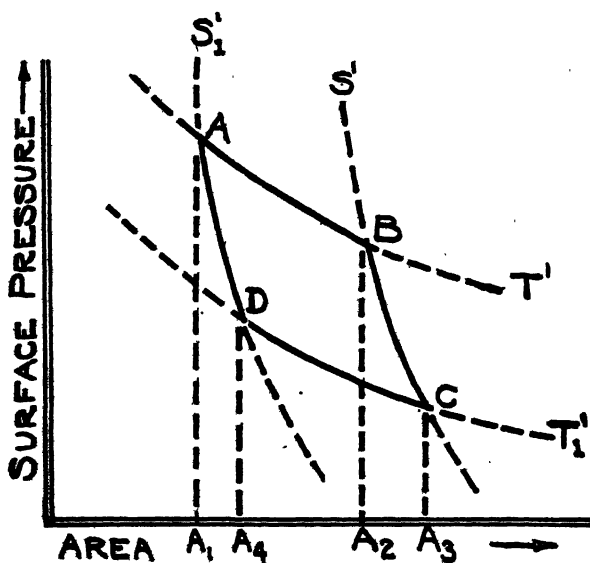
FIG. 9. THE COLLOID ANALOG OF CARNOT'S CYCLE

In figure 9 the coordinates represent the pressure and area of the surface solution of oleate we have just considered.   The curves S', S_1', represent adiabatic changes, the film being compressed or expanded without addition of energy.   It is obvious that in expanding the film and doing work, the potential, T', (or pH), of the oleate molecule will decrease, just as a gas cools when it expands doing work against outside pressure.   Any one can convince oneself of this fact by blowing air (free from CO_2) through a soap solution, collecting the foam and comparing its pH with that of the original solution.   Miss E. Laing, has measured the increase in acidity quantitatively (20).

The curves T' and T_1', are lines of equal potential, corresponding to isothermal changes in gases. We now carry the film through the following cycle: (a) expansion of surface from A_1 to A_2 at constant potential (with addition of alkali); (b) adiabatic expansion (without addition of alkali) from A_2 to A_3; (c) compression from A_3 to A_4 at constant potential (removal of alkali); (d) adiabatic compression from A_4 to A_1, back to the

Note: p_H determined by indicators.

FIG. 10. pH OF 0.3 PER CENT SOLUTIONS CONTAINING VARIOUS PROPORTIONS OF NA Ol AND H Ol

original state. The work done in this cycle is equal to the area $ABCD$ the energy being derived from the transferrence of a certain amount of matter from a potential T' to a lower potential T_1'. By proceeding as in the case of Carnot's cycle, it can be shown that the maximum efficiency of the above engine is

$$\frac{T' - T'_1}{T'}.$$

That this formula agrees with the facts, at least qualitatively,

will be evident from figures 10 and 11. Figure 10 shows the pH of a 0.3 per cent oleate solution, composed of different proportions of sodium oleate and oleic acid; Figure 11 represents roughly the variation of P', the surface pressure of the oleate solution, with change in pH. It can be seen on comparing the two curves that by adding to an oleate solution of pH 5.5, one-fifth of its equivalent of sodium hydroxide, the pH will rise to 7.5 and the surface pressure from 21 to 34 dynes/cm. The same amount of alkali

Fig. 11. CHANGE OF SURFACE PRESSURE WITH pH

(Analogous to change of gas pressure with temperature)

added to an oleate solution of pH 8.0 will increase its pH to 8.8, corresponding to an increase in surface pressure of 3.5 dynes cm. only. Since equal quantities of alkali added to the oleate solution at pH 5.5, and pH 8.0 represent (very nearly) equal amounts of energy, and since the pressure developed is a measure of the work obtainable, it follows that, other things being equal, the efficiency of the colloid engine is greater when it works at low potential.

Colloid energy E'. The energy which causes the expansion and the pressure in the colloid system considered above is derived, in the light of accepted terminology, from the chemical reaction NaOH + oleic acid. We may leave it to linguistic experts to decide whether or not there are any grounds for speaking about a "chemical change" under colloid conditions, when no crystalline compound is present, when the "combination" between NaOH and oleic acid takes place in all proportions, and the change in properties is a gradual one. It is sufficient to point out that Krafft (21) has maintained that in soap solutions oleate radicles still exist as oleic acid, basing his opinion on the identity of the solidification points of soap sols and of the corresponding fatty acids. The present writer, as already mentioned, would prefer to describe colloids as *"physical compounds,"* seeing that, in the case of surface films, for instance, the amount of sodium "combined" with the oleate radicle has been calculated by means of Gibbs' adsorption formula (22).

For our purpose all we need to bear in mind is that colloid systems store their energy not as heat energy, but as electrochemical energy; that a change in colloid volume or pressure will be accompanied *not by a thermal effect,* but by a "chemical" or electrical change. It has already been mentioned that when a soap film is expanded, it becomes more acid. Lord Kelvin (then William Thompson) calculated in 1859 that a soap film cools when it expands (23). Since the system is not "ideally" colloid a change in temperature may well occur, but if so, it is probably of secondary importance. Even when the increase in surface takes place in pure liquids, as in the spraying of water, electrical changes have been observed and measured (24). Again, when an ordinary insulating tape, which contains colloid material, is quickly unwound in the dark, there is no appreciable heat effect, but a distinct fluorescent light is emitted, due no doubt to the rapid extension and contraction of the colloid. Many other examples could be given showing the occurrence of electrical and chemical phenomena caused by sudden changes in the colloid variables P' and V'.

Conversely, the electrification of a soap or saponin bubble

causes an increase in its surface (24a), while the effect of temperature changes is either insignificant or causes the bubble to break. A potential difference across an interface leads to emulsification (24b) and chemical energy is transformed into surface energy in spontaneous dispersion.

Colloid Entropy S'. Following the example of thermodynamics the colloid potential T' has been defined by means of the ideal colloid, and also as a work function. Colloid entropy, S', like its thermal equivalent, may be defined by the integral $\int \frac{dE'}{T'}$. Its meaning may be gathered from the following material in quotation marks. It is a copy of a discussion of entropy to be found in Lewis and Randall's Thermodynamics (1923, p. 114) except that a soap film and two "reservoir solutions" at different pH have been substituted for Lewis and Randall's metallic spring and their two reservoirs at different temperatures.

"We will choose a standard system composed of a soap film and a reservoir of colloid energy. In employing this film-reservoir in conjunction with other systems, we are going to use the film as a source of work and the reservoir as a source or sink of colloid (electro-chemical) energy. It would be desirable to choose them so that the film will undergo no change in colloid potential (pH), and the reservoir will do no work during the processes we are about to consider.

"If the film is released and by some process gives up a part of its energy to the reservoir in the form of colloid energy, (e.g., transformation of oleic acid into sodium oleate, corresponding to an increase in pH due to contraction), we might measure the extent of this irreversible process by a pointer and scale attached to the film, or by the amount of energy given to the reservoir. We shall in fact take as the measure of the extent of this standard universal process a quantity which is proportional to the energy exchange, but not equal to it, for it is necessary to our purpose to consider also the colloid potential of the reservoir.

"To make this clear, we may consider a film, and two separate reservoirs, one at the colloid potential, (pH), T_1', and one at the lower colloid potential T_2'. If the film be released and a certain

amount of colloid energy is given to the reservoir at T_1', and if then this same amount of energy is allowed to flow to the other reservoir at T_2', this latter is also an irreversible process. The net result is the same as if the energy developed by the film were given at once to the reservoir at lower pH. Now the sum of the degradation in two successive irreversible processes must be greater than that in either one alone; otherwise our definition would not be quantitative. Therefore, if we are to have a genuine scale of irreversibility, the transfer of energy from the film to the reservoir at higher pH must be regarded as a less irreversible process than the transfer of the same amount of energy from the film to the reservoir at lower colloid potential (pH).

"It will therefore be expedient to define the extent of irreversibility of our standard process by making it equal not to q, but to q'/θ', where q' is the energy transferred and q'/θ' is some quantity which quantitatively satisfies our definition of colloid potential. Moreover when the function is determined, it completes the quantitative definition of degradation. Let us consider a system composed of an ideal colloid in which, by definition

$$\left(\frac{\delta E'}{\delta V'}\right)_{T'} = 0$$

For such a colloid

$$q' = \theta'\left(\frac{\delta P'}{\delta \theta'}\right)_{V'}$$

and we see that at constant volume the pressure is proportional to T', hence θ' is proportional to T'. This is all that we need to know in order to permit the complete identification of the "colloid-dynamic" scale with the ideal colloid scale."

In this particular system the transfer of energy is bound up with the transfer of ions, but we could construct a system involving a transference of electrons only, for instance involving an increase in the surface of a soap film on electrification. In that system no "matter" in the ordinary sense would pass from the higher to the lower colloid potential.

Some people may object to the introduction of several kinds of entropy, but Swinburne (25) has pointed out long ago that

ordinary thermodynamic entropy is made up of a number of different entropies. When a perfect gas expands in vacuum its entropy increases and it is only by a stretch of imagination that we conceive this as an increase in *heat* entropy. Since we speak about several types of potential, we must also assume several entropies. The second law of thermodynamics, in the light of this extended entropy principle, is a special case of a more general statement, namely: *it is impossible to obtain work by transferring energy from a lower to a higher potential*, and is merely the negative of the definition of potential which states that "it is possible to obtain work by transferring energy from a higher to a lower potential."

Muscular action. The colloid engine enables us to understand the principle of muscular action. It has often been pointed out, that the contraction of a muscle is in some way connected with "surface" forces. Many theories have been put forward (26), but none in a form suitable for exact formulation. Galeotti's interpretation of the contraction mechanism as due to changes in pH is especially interesting. According to this author (27), "the anabolic phenomenon might consist in the formation of an organic acid within the contractile elements; the catabolic phenomenon in the dissociation of this acid and in the migration of the H-ions outside the contractile elements; with this mechanism the energy accumulated during the integrative period would be transformed into work. The H-ions, diffusing through the contractile elements and combining with the OH-ion of the sarcoplasma, would produce as heat of neutralization the heat which appears during the contraction of the muscle."

If this were true the analogy between the muscle and the *special type* of colloid engine working between different H-ion concentrations would be complete, but the energy changes in the active muscle, as Galeotti himself points out, must be more complex. It is known that lactic acid is formed during contraction and that surface films of muscle protein contract when the H-ion concentration increases. Although Gorter and Grendel (28) have found that surface films of muscle protein expand when alkaline and contract when acid, Verzar (29) claims that in order to

produce artificial muscular contraction the H-ion concentration must be increased to pH 3, a much higher acidity than can be detected in the normal muscle. This objection, according to Meyerhoff, (30), is not valid, the production of lactic acid in the muscle being localized so that the reaction of the muscle as a whole need not be materially affected. The accurate work of Hill and Meyerhoff (30) shows, however, that the energy changes in the muscle are derived from the transformation glycogen $\rightleftharpoons$ lactate, and not merely from the reaction $H + OH$ (lactic acid $\rightleftharpoons$ lactate) which supplies but a fraction of the total energy.

It should be clearly realized that, for the understanding of the dynamic principle of muscular activity, we need not bring into discussion the chemical changes which supply the energy, any more than we should have to know what fuel is burnt in a locomotive in order to understand the expansion of steam and the movement of the piston. A change in pH is not essential, either; a reversible oxidation-reduction reaction (e.g. glycogen $\rightleftharpoons$ lactate) can equally well serve as a source of energy. In this case the H-ion concentration would still be of great importance, because the value of an oxidation-reduction potential depends greatly on pH (31). Furthermore, the working of the muscle is not dependent upon "surface" tension or upon the existence of a monomolecular layer. All that need be postulated for a mechanistic explanation of muscular action is an oriented arrangement of polar molecules, the distance between which is regulated by the potential T'.

THE ADSORPTION EQUATIONS

In the light of the new colloid variables P', V', T', the adsorption equation assumes a simpler and more general form which is, in fact, nothing more than the equilibrium equation.

$$V \, dP = V' dP' \tag{i}$$

V, V' and P, P' being the specific volumes and pressures of a given substance in two different regions. This general relation tells us, for instance, how the vapor pressure of a liquid varies with the external pressure, or with the osmotic pressure of the liquid,

or with the swelling pressure of a gel, and many other things (see section on generalized thermodynamics). The relation (i) may be expressed in the form

$$\frac{dP'}{dP} = \frac{V}{V'} = \frac{C'}{C} \tag{ii}$$

C' and C being the concentrations of the adsorbed substance in the surface region and in the bulk of the solution respectively.

Gibbs' equation (32)

$$\Gamma = -C\frac{d\sigma}{dP} \tag{iii}$$

is readily obtained from (ii) because

$$\frac{\Gamma}{\delta} = C'; \frac{F}{\delta} = P' \text{ and } dF = -d\sigma$$

where δ is the thickness of the surface region.

There are several reasons for preferring equation (ii) to (iii):

(a) Gibbs' equation deals with concentration per square centimeter and with tension per line, while equation (ii) refers to concentration per cc. and to pressure per area, which are more readily visualized.

(b) Gibbs' equation does not tell us what happens after the surface tension has reached its lowest value and we continue to increase the concentration of the solution; equation (ii) leads us to believe that the thickness of the surface region will change under those conditions.

(c) Gibbs' equation leads to impossible results when the σ — C curve has one or more minima, but the new equation can be adapted to such cases (see the effect of latent energy, λ, below). Equation (iii) is generally combined with the van't Hoff or "gas" law for dilute solutions $PV = RTM$ or $dP = RTdC$ so that

$$\Gamma = -\frac{C}{RT}\frac{d\sigma}{dC} \qquad \Gamma = -\frac{d\sigma}{RT\,d\ln C} \tag{iv}$$

Gibbs himself never used the adsorption equation in this form (33).

The exponential equations. If we assume the "gas law" at the surface region as well as in solution then

$$dP = RT \, dC \text{ and } dP' = R'T'dC' \tag{v}$$

Substituting these values of dP and dP' in equation (ii), we have

$$R'T' \frac{dC'}{C'} = RT \frac{dC}{C} \quad \text{or} \quad R'T' \, d\ln C' = RT \, d\ln C \tag{vi}$$

which on integration gives

$$R'T'\ln C' = RT \ln C + \text{constant} \tag{vii}$$

If we write λ for the constant of integration, we have,

$$C' = C^{\frac{RT}{R'T'}} e^{\frac{\lambda}{R'T'}} \tag{viii}$$

λ being the energy necessary to transfer one mol of the adsorbed substance to the surface region; it is a constant at constant T', but may have different values according to the configuration of the adsorbed molecules. This, together with the variation of T' on dilution, could account for the several minima observed in $\sigma - C$ curves (34).

Introducing the activity coefficient, α, which corrects for divergence from the gas law, we obtain

$$C' = \frac{\alpha}{\alpha'} C^{\frac{RT}{R'T'}} e^{\frac{\lambda}{R'T'}} \tag{ix}$$

a more general equation than (viii).[3]

Boltzmann's distribution law

$$C' = \frac{p_1}{p_2} Ce^{\frac{\lambda}{RT}} \tag{x}$$

has been applied by Langmuir (35) to the distribution of molecules between phases and interfaces, p_1 and p_2 being defined as the *a priori probabilities* of the molecules in the two states under

---

[3] An even better formula would be one taking into consideration the influence of curvature.

consideration. Formula (x) becomes identical with (ix) when $R'T' = RT$ and $\frac{\alpha}{\alpha'} = \frac{p_1}{p_2}$.

The adsorption equation derived from Boltzmann's distribution law is free from the objections raised to the Gibbs' equation. The new formula (VIII) is preferable to the Boltzmann formula because it tells us how the concentration at the surface changes with the colloid potential (or pH in the case of fatty acids). This factor *must* be taken into consideration, for when a solution of sodium oleate is diluted beyond a certain point its alkalinity decreases (pH falls), and this greatly affects the adsorption. Any equation which ignores this change in pH cannot adequately represent the effect of dilution on surface concentration. Nor can it be maintained that the fatty acids are peculiar in this respect: surface active substances are, *par excellence*, hydrolyzable compounds and the pH of their aqueous solutions changes with dilution. Even if the pH were not a variable, e.g. in nonaqueous solutions, some other factor representing the colloid potential would have to be taken into account.

Freundlich's equation

$$C' = a\, C^{\frac{1}{n}}$$

can be obtained from (viii) by assuming $R'T'/RT = $ constant $ = n$, and $e^{\frac{\lambda}{R'T'}} = $ constant $ = a$. We may also mention here *J. J. Thomson's formula* (36) derived by means of generalized dynamics, using the Hamiltonian and Lagrangian functions. It has since been deduced thermodynamically by *Michaelis* (37) and has the form

$$C' = C\, e^{-\frac{1}{\delta RT}\frac{d\sigma}{dC}}$$

But

$$-\frac{1}{\delta}\frac{d\sigma}{dC'} = \frac{dP'}{dC'} = R'T'$$

and the formula reduces to

$$C' = C\, e^{\frac{RT}{R'T'}}$$

In the light of (viii) this equation obviously is incomplete, for (viii) would require $R'T'$ to equal RT, and RT to equal λ at the same time.

The "gas law" in colloid systems. Conflicting opinions have been expressed recently concerning the validity of the relation

$$\text{force} \times \text{area} = \text{constant}$$

for surface films. Marcelin (38) found this to hold true for oleic acid films; Delaplace (39), working with benzyl benzoate films even extended the formula to $FA = KT$. N. K. Adam (40) confirms Langmuir's deduction that $FA = RT$ when the concentration of the substance at the surface is very dilute, but maintains that the law fails even at moderate concentrations.

From the point of view of colloid dynamics, very dilute films in which the molecules are free to vibrate and possess heat motion are of little interest; but the question of the behaviour of saturated or nearly saturated surface solutions is of great importance. It should be pointed out, however, that for the particular purpose of testing the gas laws in colloids, a film of fatty acid is a very unsuitable system; it corresponds to a gas at very low temperature, and under such conditions even permanent gases are apt to present anomalies. It has been shown, in fact, by Windisch and Dietrich (41) that dilute solutions of fatty acids may be surface active or not, according to conditions. This may account for the difference in the findings of N. K. Adam and of Marcelin.

In order to obtain definite proof whether the ideal gas law can be applied to colloids or not, it is necessary to work with a system at high colloid potential: for instance, sodium oleate, instead of oleic acid. The gas law states that the pressure of the system is directly proportional to (a) the concentration, and (b) the temperature or potential. The following facts are indications that in many instances this law applies (approximately) to surface solutions and gels.

(i) Freundlich (42) shows that by differentiating his equation $F = AC^{\frac{1}{n}}$ with respect to C, we obtain

$$-\frac{d\sigma}{dC} = \frac{A}{n} C^{\frac{1}{n}-1}$$
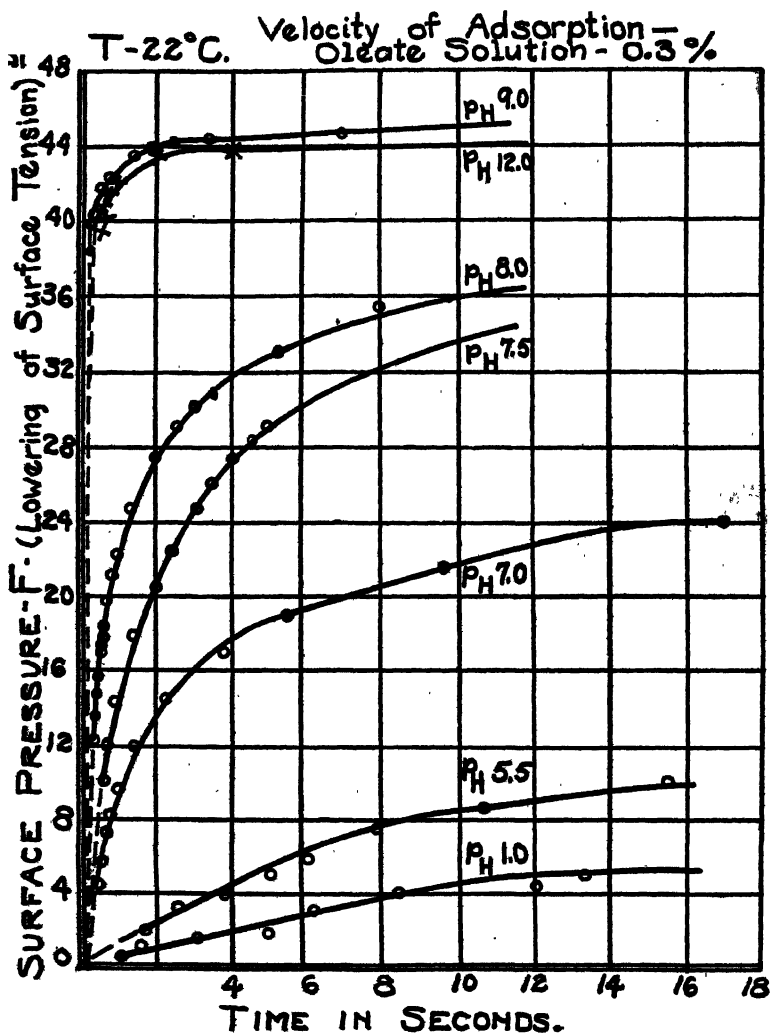
FIG. 12. VELOCITY OF ADSORPTION OF A 0.3 PER CENT OLEATE SOLUTION AT 22°C.

This, when substituted in the so-called Gibbs' equation (iv), gives $C' = KC^{\frac{1}{n}}$. In other words the surface pressure F must be proportional to the surface concentration C'.

(ii) The change in the surface pressure of sodium oleate with time can be expressed by the formula

$$\frac{F}{t \; F_\infty - F} = \text{const.} \tag{xi}$$

F_∞ being the maximum effect on the surface.[4] Differentiating with respect to time we have

$$\frac{dF}{dt} = k \, (1 - F)^2$$

[4] Johlin (44) has investigated this phenomenon quantitatively by means of capillary rise measurements and has proposed a formula of the type $\sigma = \dfrac{a}{t^2}$, in which a is a constant. His results, when experimental conditions were such as to allow of accurate determinations, fit well with formula (xi). The capillary rise method is not quite suitable for adsorption-time measurements because at usual concentration (0.1 per cent) most of the total surface tension lowering takes place within the first few seconds, and it is impossible to take readings in that time; with dilute solutions (0.001 per cent) on the other hand, long exposures to air are apt to modify the surface film.

The curves shown in figure 12 were obtained by Cofman and Sheely (unpublished data) by the "pressure in bubble" method. The apparatus used was the same as in figure 6, only in place of the bent tube used to measure the foam, a short vertical tube was substituted. The glass capillary had a radius of about $\frac{1}{200}$ cm. The air pressure was varied and the time between successive bubbles recorded.

The relation between radius of capillary, pressure, and surface tension of solution has been expressed in different ways (45). One of these is (46).

$$\frac{r}{2} P \left(1 - \frac{2}{3} \frac{\rho \, r}{P} - \frac{1}{6} \frac{\rho^2 \, r^2}{P^2} \right)$$

ρ being the density of the solution.

Since in our experiments r = 0.0025 cm. and p was never less than 100 dynes/cm. the correcting factors containing powers of r are insignificant and the formula reduces to

$$\sigma = \frac{r}{2} P$$

The sodium oleate solutions at lower pH were obtained by adding dilute hydrochloric acid. The curves with these solutions were not always exactly reproducible (see previous reference to surface active and surface inactive fatty acids).

Liepatoff (43) has found for the rate of change of adsorbed *mass* with time, for various substances, a general relation

$$\frac{dM}{dt} = k \, (a - \gamma \, M)^2,$$

γ being practically unity. This also supports the proportionality between concentration of adsorbed substance and colloid pressure.

(iii) The swelling pressure experiments bring further evidence pointing in the same direction.

(iv) If we measure the colloid potential in terms of pH, it appears to be proportional, within limits, to P', as indicated in Figure 11, provided the number of oleate molecules in a saturated film does not vary greatly with the pH.

All one can say at present, is that there exists some evidence in favor of the view that certain colloid systems are governed (approximately) by the law $P'V' = R'T'$. The ideal colloid, like the perfect gas, is non-existent.

CORRELATION OF COLLOID PROCESSES WITH ORDINARY PHENOMENA

Having shown that the new ideas may be used to derive exact laws which rule colloid systems, we shall proceed to discuss briefly a few other colloid phenomena and their equivalent crystalloid or thermal processes. It will be shown that the new concepts throw an interesting light on many obscure points, and open up new avenues of approach to many fields of inquiry.

Nerve conduction. This is not the place for discussing at length the physiological aspects of the new theory of colloids, but it may be pointed out that from the relations between the variables P', V', T', etc. it follows that a compression, or a chemical change, or other disturbance at a point in a colloid will cause a variation in pressure and potential. As Michaud (47) remarks: "one can hardly touch a gel without causing a difference in electric potential" which can be readily detected with a galvanometer. A disturbance produced in any manner in a colloid will propagate

itself with a velocity depending on the elasticity and density of the medium. Just as in the case of propagation of sound-waves

$$\text{Velocity} = \sqrt{\frac{\text{Elasticity}}{\text{Density}}} \qquad \text{(xii)}$$

If the colloid is "ideal" and obeys the gas law $P'V' = R'T'$ we obtain by differentiating at constant T'

$$P' = - V' \frac{dP'}{dV'} = \text{Elasticity} \qquad \text{(xiii)}$$

Remembering that in colloids we must use surface instead of volume, we have

U, the area per unit mass, instead of 1/density.

Therefore

$$\text{Velocity of compression waves} = \sqrt{P'U}.$$

Both P' and U can be determined directly from measurements on surface films. No such determinations have yet been made apparently for the protein matter of the nerves, but for the proteins of the rabbit serum we have, from the data of du Noüy (34) $P' = 17$ dynes/cm. $U = 1.8 \times 10^6$ cm.^{2} Therefore

$$\text{Velocity} = \sqrt{17 \times 1.8 \times 10^6} = 56 \text{ m/sec. approximately.}$$

It is remarkable that the velocity thus calculated is of the same order of magnitude as the velocity of propagation of nerve stimuli in warm-blooded animals. This good agreement is no doubt largely fortuituous, there being large uncertainties in our calculation: (a) The values of P' and U for the protein matter of the nerve may be different from those of the serum, though Gorter and Grendel (28) have found the same value for U in the case of muscle protein; (b) the elasticity in formula (xiii) refers to conditions of constant potential. The wave of compression is probably an adiabatic phenomenon and a factor must be introduced in the calculation to take care of this, but this will not affect the order of magnitude of the velocity; (c) the colloid potential T' must be taken into consideration.

A critical survey of the literature on nerve conduction would fill a volume by itself. The reader may turn for information to Cremer (48) and Hallowell Davis' (49) recent review.. It may be mentioned here that it is generally accepted by physiologists (50) that the nerve stimulus is propagated as "a molecular disturbance, wave-like in character.[5] Erlanger, Gasser and Bishop (51) by means of their cathode ray oscillograph have photographed the wave and shown it to be a potential wave. Sutherland (52) proposed the same formula (xii) for the velocity of transmission of stimuli, calculating the elasticity from Young's modulus for gels. Broemser (53) starting from quite different premises, reaches a similar formula in which the osmotic pressure of the liquid surrounding the nerve takes the place of the elasticity or the swelling pressure of gel. But we have seen that the osmotic pressure is identical with the pressure P' if the colloid is impermeable to the crystalloid constituents of the solution, which is assumed to be the case in nerves. Consequently the experimental support which Broemser brings for his formula serves to further strengthen formula (xii).

The semipermeable membrane. Theoretically, a semipermeable membrane is a thin wall separating two phases, allowing certain kinds of matter to diffuse through and not allowing others. Practically, a semipermeable membrane is a more or less ideal colloid (gel) or a rigid system having large surface, and therefore, possessing colloid properties.

Consider a gelatine gel such as was used in the swelling pressure experiments (fig. 3). If one side of the gel is in contact with a sugar solution and the other side with pure water then the gelatine acts as a semipermeable membrane.

---

[5] The "all or none" principle which states that the stimulus traveling along a nerve has always the same intensity, is sometimes assumed to contradict the wave theory of transmission. In reality it is not incompatible with it, but requires additional hypotheses to account for the supply of energy on the way. A careful survey of the evidence in favor of the "all or none" principle shows that it is far from being completely verified. Many of the experiments which were originally supposed to prove this principle were shown by subsequent investigators to have been wrongly interpreted (54). At the present time the only support for the "all or none" principle, to the exclusion of all former "proofs" is that supplied by the experiments of Davis and his co-workers (54).

The immediate mechanism of semipermeability is here easily understood. The gelatin next to the sugar solution loses more water than that in contact with pure water. (Fig. 13). The swelling pressure in the membrane is therefore greater at A then at B and water flows from B to A. In consequence, water passes from the pure liquid to the solution. At equilibrium the sugar solution will be at a uniform pressure P_1, the water at another uniform pressure P_2. In the colloid region (membrane) the pressure will vary from P_1' at A, to P_2' at B. It is the rigidity of the colloid membrane allowing such a difference of pressure to exist in the colloid region which makes possible the measurement of osmotic pressure. What is measured in osmotic cells is not the osmotic pressure, but this difference in colloid pressure on the
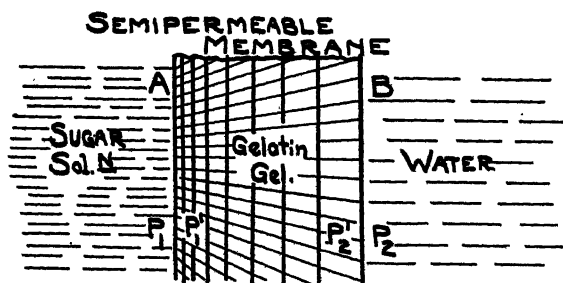
FIG. 13. ACTION OF A GELATINE GEL AS A SEMI-PERMEABLE MEMBRANE

two sides of the membrane. Only when none of the solute particles pass through, which is seldom the case, is the difference in colloid pressure at A and B equal to the osmotic pressure. In the above experiment, where the semipermeable membrane is an elastic gel, the difference in pressure on its opposite sides is brought about by a change in the concentration of the colloid. If the semipermeable membrane is rigid, (e.g., ferrocyanide membrane) no change in concentration can occur The colloid potential T' or some other factor on which colloid pressure depends must then be different at A and B, in order to produce the necessary difference in pressure.

Electro-osmosis and "thermo-osmosis." An obvious instance, where the difference in pressure on the opposite sides of the semi-

permeable membrane is caused by a difference in colloid potential, is electro-osmosis.

We have seen that the colloid potential depends on the electrochemical potential; an applied e.m.f. will, as a rule, produce a difference in colloid potential. This in turn will cause a difference in pressure on the two sides of the colloid region and liquid will flow across the membrane.

In order to fix our ideas, let the semipermeable membrane consist of a gel of sodium oleate, firmly fixed in a tube so that its volume cannot change (fig. 14). If a sufficiently high e.m.f. is applied, alkali will be liberated on that side of the membrane

ELECTRO-OSMOSIS

"OSMOTIC FLOW"
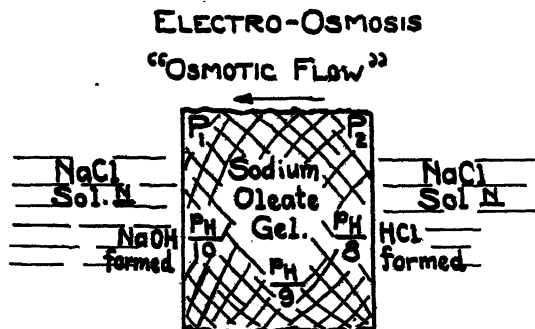
FIG. 14. ELECTRO-OSMOSIS

facing the positive pole and acid on the opposite side (55). The swelling pressure P' will increase on the alkaline side (higher colloid potential) and will decrease on the acid side, thus producing the necessary conditions for osmotic flow.[6]

An analogy with a crystalloid system will make the matter clear. Suppose that a sugar solution is divided by a semipermeable membrane into two compartments, maintained at different temperatuess. The osmotic pressure will be greater on the warmer side ($P \infty T$) and pure solvent will flow from the cold to

---

[6] In a gelatine gel water will flow from the positive to the negative side or viceversa, according as the gelatine is on the acid or on the alkaline side of the pH.

the warm section.   This is clearly the thermodynamic equivalent of electro-osmosis and might be termed thermo-osmosis.[7]

Stream potential and "stream temperature."   If we reverse the procedure in the above experiment, that is, if the temperature in the two compartments is equal to begin with, and pure solvent is caused to flow across the semipermeable membrane by applying pressure on one side, then a difference in temperature will develop. This is accounted for in terms of "heat of dilution": we may call it "stream temperature," to conform with the name applied to the same phenomenon in colloid systems: the difference in colloid potential produced on the two sides of the membrane, measured as an electrical potential difference is called "stream potential."

Electrophoresis and "thermo-phoresis."   When colloid particles are placed in a non-uniform electric field, they acquire, in general, a different colloid potential at front and rear and move according to circumstances, either with or against the impressed electric field.   For reasons given in the paragraph on "imperfect colloids," it is not easy to give a simple colloid-dynamic interpretation of this phenomenon.   However, even here it is possible to draw attention to an equivalent thermodynamic phenomenon:

A light body whose opposite sides are at different temperatures will move one way or the other.   This happens for instance to the vanes in Crook's radiometer.   The conditions in the two instances are not identical, but it is possible to construct an even closer analogy: Consider a small amount of sugar solution enclosed in a semi-permeable membrane.   A temperature gradient in the surrounding liquid will set such a system in motion because, as in thermo-osmosis, water will penetrate at one end and be eliminated at the other.

The opposite phenomenon, namely the existence of a difference in temperature between the front and rear of an object moving in a gas or liquid, can hardly be doubted.   The movement of a

---

[7] In Soret's experiments—in the absence of a semipermeable membrane—a difference in temperature produced differences in the concentration of the solute present in solution.   [Ch. Soret: Ann. Chim. Phys., **22**, 293 (1881); J. van't Hoff: Zt. Phys. Chem., **1**, 487 (1887)].

body produces compression ahead and rarefaction behind, and this must give rise to opposite temperature effects.

The colloid dynamic phenomenon that corresponds to the last mentioned thermodynamic effect is the potential difference due to falling particles.

The imperfect colloids. The new variables have not yet been applied to the subject of lyophobic colloids (such as gold sols). These systems may be described as "imperfect" colloids and are subject to more complicated changes. They can be compared with vapors as contrasted to permanent gases. A vapor, in a vacuous enclosure provided with a movable side, is an extremely unstable system: a slight change in temperature or pressure may cause the formation of a mist, or complete liquefaction (large decrease in volume).

Lyophobic colloids similarly are very sensitive to slight changes in the conditions of the dispersing medium. A trace of electrolyte may cause partial or complete precipitation of a suspension (large decrease in surface, or colloid volume).

The water-vapor analogy can be carried yet further: The presence of a permanent gas in an enclosure "stabilizes" the vapor. For instance, consider a cylinder provided with a movable piston, and which contains air mixed with water vapor in contact with liquid water. It is possible for the cylinder to contain water vapor under conditions of temperature and pressure which would condense pure water vapor. In the same way, a lyophilic (permanent) colloid may stabilize a lyophobic system.

The laws which define imperfect colloid systems are naturally more complex than those which apply to permanent or ideal colloids. It is perhaps because most of the work on colloids has dealt with those difficult systems, that the simple laws of colloid dynamics have for so long escaped detection.

One other factor complicates lyophobic systems, namely, the curvature of the surface, which apparently has an influence similar to that of gravitation in ordinary systems. For instance, any particle floating on the surface of water in a beaker is "attracted" towards the greater curvature near the edge, if only it approaches close enough to the sides of the beaker. The con-

centration of an adsorbed substance is also greater in the curved region. This is indicated by the following experiment which incidentally demonstrates the reversibility of the adsorption phenomenon. If a little foam (obtained by blowing air through a capillary into sodium oleate solution) is placed on the surface of pure water, it breaks immediately; if placed on soap solution of concentration greater than 0.02 per cent (at pH 10) the foam lasts a long time, but when the soap concentration is decreased, a point is reached when the bubbles on the flat portion of the surface break immediately while those near the edge persist.

Many of the considerations in the last sections are of a qualitative nature only, and must await quantitative confirmation. In view of their simplicity and symmetry the writer is confident that quantitative proof will be forthcoming and possibly in some cases can be obtained from existing data. It is pertinent to remark that one of the most widely quoted surface "laws," the Gibbs' equation, has had so far only semi-quantitative confirmation (33).

Extensive use has been made above of the analogy between colloid and crystalloid or thermal systems. Analogies are often deceptive and the fact that in this instance it has been possible to carry them so far without coming across obvious incompatibilities, leads one to suspect that there is some "natural law" which causes the variables P' V' T' to behave in the same manner as P V T etc. This point of view will be further enlarged in the next section on generalized thermodynamics.

GENERALIZED THERMODYNAMICS

Consider Gibbs' fundamental equation (56)

$$dE = TdS - PdV + F_n dM_n \qquad (i)$$

M_n and F_n stand for different kinds of substances and their corresponding "partial molal free energies." Although dE has no subscript it also stands for various kinds of energy; similarly P and V may represent either the pressure and volume of gas, or osmotic pressure and its corresponding volume. We have seen,

too, that there are several T's and S's therefore we may write formula (i) more consistently

$$dE_n = T_n dS_n - P_n dV_n + F_n dM_n \qquad (ii)$$

where n stands for any number of variables of the same type For instance M_n may represent the mass of gas, colloid, electrons, etc.; P_n may refer to gas pressure, colloid pressure, osmotic pressure, electromotive force, and so on.

The practical value of this scheme is best shown by an example: We may express a large number of apparently unrelated physical "laws" by simply combining the gas law $PV = RT$ with the general equilibrium condition

$$P_x dV_x = P_y dV_y$$

where x and y refer to different types of systems (electronic, molecular, macroscopic). For instance:

Vapor pressure and total pressure on the liquid

$$\frac{dP_v}{dP_l} = \frac{V_l}{V_v} \qquad \text{or} \qquad (VdP)_l = (VdP)_v$$

Vapor pressure and osmotic pressure. The osmotic pressure P_o corresponds to a decrease in internal pressure of liquid P_l; therefore

$$- dP_o = dP_l$$

Now,

$$- P_o = \frac{RT}{V_l} \ln P_v + \text{const.} \qquad \text{or} \qquad - dP_o = dP_l = \frac{RT}{V_l} \frac{dP_v}{P_v}$$

hence

$$(VdP)_l = (VdP)_v$$

Vapor pressure and swelling pressure of gels. The formula is identical with that for osmotic pressure (57).

The adsorption formula

$$\Gamma = - \frac{C}{RT} \frac{d\sigma}{dC}$$

has already been shown to be a special case of $(VdP)_1 = (VdP)$ surface.

Nernst's equation for electromotive force. In a cell with a hydrogen electrode the relation between e.m.f. and gas pressure is

$$E = E_o + \frac{RT}{Fr} \, lnp \quad \text{or} \quad Fr.dE = \frac{RT}{P} \, dP \quad \text{or} \quad (Fr.dE) = (VdP)_{gas}$$

Here Fr. stands for one Faraday of electricity and, in this formula, it is apparently a volume and not a quantity factor.

Atmospheric pressure and altitude (h)

$$\frac{1}{P}\frac{dP}{dh} = -\frac{g}{RT} \quad \text{or} \quad -VdP = gdh \quad \text{or} \quad (PdV) = (gdh)$$

In the same way the Clapeyron equation, the Gibbs-Helmholtz equation for the temperature coefficient of e.m.f., the Richardson thermionic equation (58), and the equation giving the latent energy of emulsification, can all be expressed by the general formula

$$\frac{\lambda}{dV_n} = T\frac{dP_n}{dT_n} - P_n$$

λ being latent energy.

No doubt the time is not distant when all physical laws will be conveniently tabulated in groups instead of being expressed in many different ways, as it is now the custom.

Formula (ii) also throws light on the 3rd law of thermodynamics. The total entropy of a system will be zero only when *all* its potentials, not temperature alone, become zero. If any kind of difference of potential exists at the absolute zero of temperature the equalization of that potential difference will cause an "increase in entropy."

Entropy and time. One of the chief drawbacks of thermodynamics has been its neglect of the time variable. This may be readily overcome by multiplying the left side of the equation (ii) with $\frac{dt}{dt}$, t standing for *time*. In the equations that follow we shall, for convenience, drop the n's. We have

$$\frac{dE}{dt} \, dt = T \, dS - P \, dV + F \, dM \qquad \text{(iii)}$$

dE/dt, which we shall denote henceforth by X, is a "power" function.   Equating (iii) to zero, we obtain

$$(F\,dM - P\,dV) + (T\,dS - X\,dt) = 0 \qquad (iv)$$

This is a *symmetrical* equation, with time as a variable.   The first thing to be noted is that the relation between entropy and time corresponds to that between mass and volume: *entropy extends in time, just as mass extends in space.   The events which compose our four-dimensional world can be separated into a space component (matter), and a time-component (entropy).*

A little reflection will show that, looked at in this way, the concept of entropy becomes more rational and easy to grasp, and fits in with all that is known about it.   Entropy is always associated with an *event*, extending in *time*.   The attempt to associate entropy with a body, or system, is responsible for the · great difficulties which still surround that subject.   The entropy of a system has been expressed in terms of the logarithm of probability.   It may be possible to express the mass of an hour of time at a given place, in terms of the logarithm of probability, but such a concept would not be a simple one.

The entropy of the world is constant.   This is a restatement of the 2nd law of thermodynamics in a new and preferable form—placing it in the same category as the first law—the conservation of energy.   Naturally, moving along the time axis we come across more and more entropy in the same way as we encounter more and more matter as we move along in space.   At a given time there are places in the world where there is no matter, and equally, at a given place there are times at which there is no entropy: the system in that space is at rest.

Living and lifeless systems.   The great value of equation (iv) lies in its simplicity and symmetry, which should prove valuable in deducing new relations among its constituent variables.   In the first place, it supplies a simple thermodynamic distinction between animate and inanimate processes.

When dealing with inanimate systems, provided we know the *usual* variables involved, it is possible to predict what will take place.   An entropy-time curve, showing the amount of entropy in each unit of time, may then be plotted in advance.

The curve AC (fig. 15) may represent, for instance, the entropy-time curve of a simple chemical reaction, and the irregular line DEFGH (or rather the area under it) may tell us the behaviour, of a more complicated system. It may be said in such instances that *entropy is fixed in time* or at least its position is determinable with our present limited knowledge. Their counterpart is a material system at equilibrium in space. But, as we know,
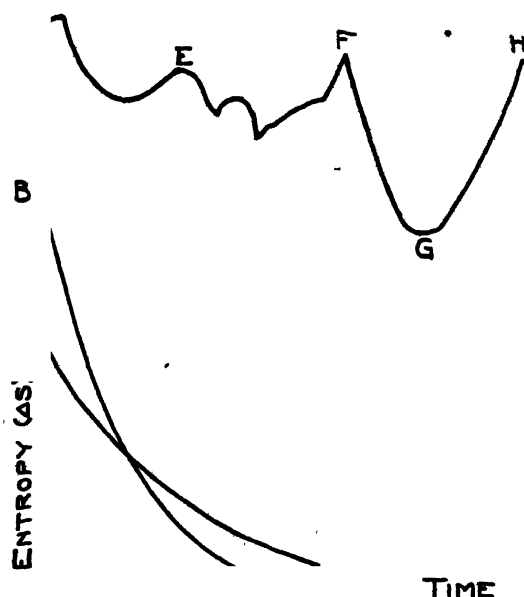
FIG. 15. ENTROPY-TIME DIAGRAM; ENTROPY FIXED IN TIME

matter can change its position in space and the question arises, what would happen if entropy changed its position in time? Obviously, if entropy were shifted along the time axis, removed from one moment and heaped on to another, then the course of events would be changed.

Consider, for instance, a watch: The gradual unwinding of its spring during a period of time, say 24 hours, is an irreversible event with which is associated a definite amount of entropy. Let

the horizontal line BKL (fig. 16) represent the distribution of
entropy in time when the watch is left by itself.   If at a moment,
K, the watch is stopped by an external agency, it comes instantly
to equilibrium and the entropy change falls to zero.   At another
moment, N, one may set the watch-spring loose, so that the whole
of its energy is liberated at once.   The entropy of the event will
at that instance jump up (NFP) and then quickly fall to zero as
the energy is dissipated into heat.
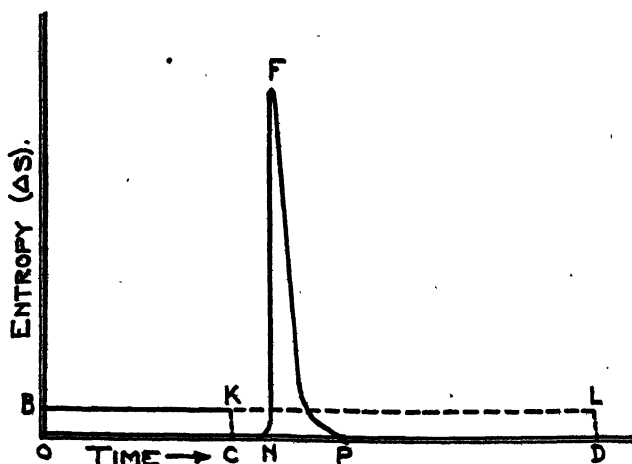
FIG. 16. ENTROPY-TIME DIAGRAM FOR A WATCH SPRING, AS AFFECTED BY
EXTERNAL FORCES

In the event just considered the distribution of entropy in time
was changed by the intervention of external forces.   Suppose,
however, that the movement of entropy were to take place in
an isolated system.   Such a system could defer its movement
and its action and behave in an entirely erratic manner, judged
from the point of view of ordinary inanimate systems, in which
entropy is fixed in time.   *It would be said to possess life.*[3]

If a catalyst were present in the system undergoing the chemical
change whose entropy is represented by area under the curve *AC*

[3] From the psychological angle, memory and foresight are in some way con-
nected with the movement of the *life function* along the time axis and should be
compatible with such a system.

(fig. 15), then the rate of reaction would be increased. The entropy-time distribution would be expressed by the steeper curve BT, instead of the curve AC. The areas under the two curves denote the total entropy of the reaction and are therefore equal; only the distribution of entropy in time is changed by the presence of catalyst. A system containing catalysts or enzymes appears to be intermediate between living and lifeless; its entropy may be considered to move in a simple manner. The increase in the rate of reaction can be taken (under given conditions) as a measure of the "activity" or mass of active catalyst present. This would be expected from the symmetry of equation (iv) since we know that (under given conditions) the change of matter in time and space is a function of the entropy (Δs).

Going one step further, it may be assumed that the more complicated shifting of entropy along the time axis, which distinguishes living systems, is also predictable in terms of a function M_x, which has the dimension of mass, but is not matter in the ordinary sense of the word. The practical aspect of this reasoning lies in the suggestion of a quantitative measure of the *life factor*. Life may be measured quantitatively by its effect on the entropy of a system. In two systems otherwise identical, but one containing life the entropy-time distribution will differ and from this difference the amount of "life" may be deduced. The subject is evidently one for experimental investigation. For simple changes which take place in the presence of both living and lifeless organic matter, e.g., yeast fermentation or liquefaction of gelatine by bacteria, some data bearing on this point should not be difficult to obtain. Even in higher animals if the nerve of one limb were severed or anaesthesized and the other not, the difference in the entropy-time distribution for the two limbs would give an indication of the amount of "life" transmitted by the intact nerve. Blood circulation would have to be prevented and other precautions taken to keep the conditions of the two limbs identical, except for the nerve impulse received by the one and not by the other. This method would not work if the reactions which supply muscular energy were of an "explosive" nature, as it has sometimes been suggested. The method

presupposes that the reactions are reversible, which is very likely, from what we know of the transformation glycogen $\rightleftarrows$ lactate.

It is not within the province of thermo-dynamics, or generalized dynamics, to give a "mechanical" interpretation of the variables involved in its formulae. Nevertheless, those who like to have a pictorial representation of abstract factors, and who wish to speculate as to the possible "nature" of the life function, M_x will find the following passage from Eddington's Mathematical Theory of Relativity, of interest. "A particle of matter is a structure whose linear extension is time-like. We might perhaps imagine an analogous structure ranged along a space-like track. That would be an attempt to picture a particle traveling with a velocity greater than that of light; but since the structure would differ fundamentally from matter as known to us, there seems no reason to think that it would be recognized by us as a particle of matter, even if its existence were possible. For a suitable chosen observer a space-like track can lie wholly in an instantaneous space. The structure would exist along a line in space at one moment; at preceding and succeeding moments, it would be non-existent. Such instantaneous intrusions must profoundly modify the continuity of evolution from past to future. In default of any evidence of these space-like particles, we shall assume that they are impossible structures (59)."

In our daily experience, is there indeed default of "structures" not recognized by us as matter, "instantaneous intrusions non-existing at preceding and succeeding moments, which profoundly modify the continuity of evolution from past to future"?

> "Bist du beschränkt, dass neues Wort dich stört?
> Willst du nur hören, was du schon gehört?
> Dich störe nichts, wie es auch weiter klinge,
> Schon längst gewohnt der wunderbarsten Dinge"
> —Goethe's "Faust" (pt. II, Dark Gallery Scene).

ACKNOWLEDGMENTS—AND AN APOLOGY

the Armour Research Laboratory. He is greatly indebted to Prof. W. P. Davey, of The Pennsylvania State College, without whose prompting this article would not have been written.

An apology is undoubtedly due to the reader. The writer feels that he has covered rather hastily a large unexplored region, and has often ventured into strange fields with many pitfalls awaiting him; in consequence, his views are probably not entirely free from errors. The only excuse he can offer for this recklessness is that he has always tried to submit his deduction to experimental proof and that the method used requires intrinsically few arbitrary assumptions; moreover, if there be serious mistakes in his deductions, the writer feels confident that, in view of the fact that he is not yet an authority, they will not be accepted as gospel truth, but will be exposed forthwith, as a warning to all rash theorizers thereafter.

REFERENCES

(1) ZSIGMONDY, R.: Kolloid-chemie, Leipzig (1920).
(2) BANCROFT, W. D.: Applied Colloid Chem., 1921 (p. 242).
(3) ALEXANDER, J.: Colloid Chemistry, Vol. 1, p. 11 (1926).
   VON WEIMARN, P. P.: Id., p. 27.
   HARKINS, W. D.: Id., p. 192.
   HARDY, W. B.: Id., p. 288.
   LANGMUIR, I.: Id., p. 525.
(4) EINSTEIN, A.: Ann. der Physik, 4, (IV), 513 (1901).
(5) POSNJAK, E.: Kolloid-Chem. Beihefte, 3, 417 (1912).
(6) BERKELEY, EARL OF, AND HARTLEY, E. G. J.: Phil. Trans., A 206, 486 (1906).
   MORSE, H. N.: The Osmotic Pressure of Aqueous Solutions, Carnegie Inst. Public., p. 105 (1914).
(7) ADAM, N. K.: This Journal, 3, 163 (1926).
(8) BOGUE, R. H.: J. Am. Chem. Soc., 44, 1343 (1922).
   McBAIN, J. W.: In Alexander's Colloid Chemistry, V. I, p. 141 (1926).
   WEISER, H. B.: In Bogue's Colloidal Behaviour, V. I, p. 377 (1924).
(9) FISHER, M. H.: Kolloid Chem. Beihefte, 15, p. 18 (1922).
(10) HARKINS, W. D., AND ZOLLMAN, H.: J. Am. Chem. Soc., 48, 69 (1926).
(11) LANGMUIR, I.: J. Am. Chem. Soc., 39, 1879 (1917).
(12) TRILLAT, J. J.: Ann de Physique, 6, p. 5 (1926).
(13) ADAM, N. K.: Proc. Roy. Soc., 101, (A), p. 519 (1922).
(14) JARISCH, A.: Biochem. Zt., 134, p. 163–175 (1922).
(15) LEWIS, G. N., AND RANDALL, M.: Thermodynamics. New York (1923), p. 63.
(16) GAD, J.: Du Bois' Archiv. f. Anat. u. Phys., p. 181 (1879).
   Also in Bancroft's "Theory of Emulsification." J. Phys. Chem., 16, 349 (1912).

(17) MADAY: Zentralbl f. Physiol., **27**, 38 (1914).
(18) FREUNDLICH, H.: Colloid and Capillary Chemistry, 1922, p. 601.
(19) HARKINS, W. D., DAVIES, E. C. H., AND CLARK, G. L.: J. Am. Chem. Soc., **39**, 541 (1917).
(20) LAING, MARY EVELYN: Proc. Royal Soc., **109**, (A) 28, (1925).
(21) KRAFFT, F.: Ber. **32**, 1596 (1899).
(22) RIDEAL, E. K.: Introduction to Surface Chemistry (1926), p. 249; from data by H. Hartridge and R. A. Peters: Proc. Roy. Soc., A **101**, 348 (1922).
(23) THOMPSON, WM.: Phil. Mag., (4), **17**, 61 (1859).
(24) NOLAN, J. J., AND GILL, H. V.: Phil. Mag., **46**, 225 (1923).
　　　LENARD, PH.: Wied. Ann., **46**, 584 (1892).
(24) VAN DER MENSBRUGGHE, G.: Pogg. Ann., **141**, 287 (1870).
　　　WILSON, C. T. R., AND TAYLOR, G. I.: Camb. Phil. Soc. Proc., **22**, 728 (1925).
(24b) COFMAN, V.: J. Phys. Chem., **29**, 1289 (1925).
(25) SWINBURNE, J.: "Entropy," 1904 (Constable & Co.) p. 107.
(26) D'ARSONVAL, A.: Archives de Physiologie, **21**, 460 (1889).
　　　IMBERT, A.: Arch. de Physiol., **29**, 289 (1897).
　　　JENSEN, P.: Pfluger's Arch., **80**, 176 (1900).
　　　BERNSTEIN, J.: Pfluger's Arch., **85**, 271 (1901); **122**, 129 (1908).
　　　VERWORM, M.: Algemeine Physiologie, 4th ed., Jena, 1903, p. 608.
　　　CREMER, M.: Zeit. fur. Biol., **47**, 562 (1906).
　　　HABER, F., AND KLEMENSIEWICZ, Z.: Zt. Phys. Chem., **67**, 285 (1909).
　　　FLETCHER, WM., AND HOPKINS, F. G.: Proc. Roy. Soc., B, **89**, 444 (1917).
　　　GARNER, W. E.: Proc. Royal Soc., B, **99**, 40 (1925).
(27) GALEOTTI, G.: Zeitsch. f. Alg. Physiol., **6**, 111 (1906).
(28) GORTER, E., AND GRENDEL, F.: Nature, **117**, p. 552 (1926).
(29) VERZAR, F.: Arch. Neerl. Physiol., **7**, p. 68 (1922).
(30) HILL, A. V.: Ergeb. der Physiolog., **22**, 299 (1923); **24**, 43 (1925).
　　　MEYERHOFF, O.: Ergeb. der Physiolog., **22**, 328 (1923).
(31) CLARK, W. M.: This journal, **2**, 127 (1925).
　　　KENDALL, E. C., AND NORD, F. F.: J. of Biol. Chem., **69**, 295 (1926).
(32) GIBBS, W.: Scientific Papers, V. 1, p. 235.
(33) LEWIS, WM. C. McC.: Phil. Mag., (6) **15**, 499 (1908).
(34) DU NOÜY, L.: 3rd Colloid Sympos. Monogr., 1925, p. 25.
(35) LANGMUIR, I.: Id., p. 48 (1925).
(36) THOMPSON, J. J.: Applications of Dynamics to Physics and Chemistry, (1888), p. 191.
(37) MICHAELIS, L.: The Dynamics of Surfaces, 1914 (Spon. Ltd), London, p. 24.
(38) MARCELIN, A.: Ann. de Physique, **4**, 459 (1925).
(39) DELAPLACE, R.: C. R., **180**, 2024 (1925).
(40) ADAM AND JESSOP: Nature, **117**, 484 (1926).
　　　COFMAN: Id., **117**, 755 (1926).
(41) WINDISCH, W., AND DIETRICH, L.: Kolloid Zt., **26**, 193 (1920).
(42) FREUNDLICH, H.: Kolloid Zt., **3**, 212 (1908).
(43) LIEPATOFF, I. S.: Kolloid Zt., **39**, 9 (1926).
(44) JOHLIN, J. M.: J. Phys. Chem., **29**, 1129 (1925).
(45) FERGUSON, A.: 5th Colloid Report Brit. Assoc. Adv. Science (1923), p. 1.
　　　SUGDEN, S.: J. C. S., **125**, 27 (1924).

(46) FREUNDLICH, H.: Colloid and Capillary Chem., (1922), p. 22.

(47) MICHAUD, F.: J. Chim. Phys., 21, 386 (1924).

(48) CREMER, M.: Nagel's Handbuch der Physiol, (IV), p. 804 (1909).

(49) DAVIS, HALLOWELL: Physiol. Reviews, 6, 547 (1926).

(50) LUCIANI, L.: Human Physiology, V. III, 3, 206 (1915).

(51) ERLANGER, GASSER AND BISHOP: Amer. J. of Physiol., 62, 496 (1922); 70, 624 (1924).

(52) SUTHERLAND, W.: Am. J. of Physiol., 17, 292 (1906): 14, 112 (1905).

(53) BROEMSER: Zt. f. Biol., 72, 37 (1920).

(54) LUCAS, K.: The Conduction of the Nervous Impulse, Longmans, 1917, p. 13.

ADRIAN, E. D.: J. of Physiol., 46, 384 (1913).

DAVIS, FORBES, BRUNSWICK AND HOPKINS: Am. J. of Physiol., 76, 470.

(55) FREUNDLICH, H.: Colloid and Capillary Chemistry, (1922), p. 264.

(56) GIBBS, W.: Scientific Papers, V. I, p. 63.

(57) KATZ, J. R.: Kolloid Chem. Beihefte, 9, 106 (1917).

(58) RICHARDSON, O. W.: Emission of Electricity from Hot Bodies, London, (1921), p. 31.

(59) EDDINGTON, A. S.: The Mathematical Theory of Relativity, Cambridge, 1924, p. 23.

· RECENT ADVANCES IN THE DETERMINATION OF THE STRUCTURE OF PROTEINS

EMIL KLARMANN

Plaut Research Laboratory of Lehn & Fink, Inc., Bloomfield, New Jersey

INTRODUCTION

In reviewing the development of the methods of research in experimental sciences and their theoretical foundations, we may observe generally that the way in which the contemporary investigator establishes or attempts to establish a scientific truth is often somewhat different from that of the scientist of several generations ago. Physics illustrates in the most distinct manner the conception of the working hypothesis now and some time ago. In considering the idea of a working hypothesis one realizes that every type of research is connected in a more or less fundamental manner with the concept of a working hypothesis. We expect from a working hypothesis that it permits a combination and correlation of facts which have been previously established in a certain realm, but at the same time it must indicate ways how and where to obtain more facts which would make the particular realm most complete and clear.

In trying to reach this goal one may encounter new experimental evidence which strengthens the foundation of the working hypothesis and raises it to the status of a theory. The latter may then serve as a firm basis from which more detailed investigation is undertaken in different directions. On the other hand one possibly meets with experimental evidence which does not fit into the general scheme of the working hypothesis; and when the disagreeing facts are placed in a sufficiently strong position by repeated confirmation, the investigator is bound to abandon the first hypothesis and must endeavor to find a new one which serves his purpose better. As a rule he attempts to

do this as quickly as possible in order to have at his disposal a guiding principle for new research work.

One is used to the idea that in experimental sciences the theory is formed in most cases in a deductive way. A great number of known phenomena is gathered and a mutual relation is discovered. This is as a rule accomplished by one individual of great intuitive ability. In most cases he is in a position to immediately secure experimental corroboration if necessary. Thus a new theory is usually based in most cases on a number of well known facts. Naturally this statement does not hold sway in all cases. It certainly does not apply to philosophy. Likewise it is not to be applied to some findings in natural sciences, particularly physics and chemistry, when these findings have been obtained in a purely speculative way. It suffices to point to the atom theory of Democritus and to the large share of scientific truth it contains, in order to realize that inductive speculation always has its justification and that intuitive investigators foresee things, the experimental corroboration of which may come centuries later. But we wish to exclude these singular instances from our considerations now and return to the scientist who builds his theories on experimental foundations and obtains the confirmation of his assumptions from experimental evidence.

And here in an ever increasing number of cases we see that the scientist proceeds not in a deductive but rather in an axiomatic manner. He does not summarize large scientific evidence. He has only little evidence at his disposal which in itself would not fill the entire volume of a theory. He therefore produces a postulate including in it the little experimental evidence he possesses and proceeding in an inductive manner he develops it into a theory and tries to devise methods and ways for the deciding experimental test. Investigators in other fields may realize that the idea contains something fundamental and they apply it, again possibly in an axiomatic manner, to their particular problems; this procedure may prove the value of the idea as a working hypothesis but it may also contribute to the strengthening of the foundation of the original theory. It is obvious that the postulate plays an extremely important part in a modern

theory.   It is sufficient to refer to Planck's quantum theory or to Einstein's theory of relativity in order to illustrate the above contention.

The reference to theories of physics in a paper on a chemical topic is made because the examples cited are particularly striking. Without dwelling upon problems of physical chemistry, we proceed immediately to organic chemistry some phases of which commence to bear resemblance to the axiomatic methods of physics.

The development of organic chemistry of natural products is a classic example of deductive work.   The scientist always attempted to obtain an analysis of the compound under investigation.   He tried to get complete information on its constituents and then to build up the same original compound with laboratory means and methods.   The synthesis was the confirmation of the analysis and the solution of the problem.   Many compounds were thus investigated and their constitution determined.   But there is still a great number of substances which are being investigated at the present time in a different manner.   Among these is the large group of proteins and their derivatives.   In order to obtain information on these compounds, scientists do much work on so called models which to our mind represent nothing else but axioms stating that a given structure is postulated for a certain chemical individual or group.   In some cases the experimental evidence for an assumed model structure is very small.   On the other hand we observe that theories are utilized which were originally devised for different realms of chemistry.   All this points to the fact that changed tactics are being adopted in organic chemistry for the elucidation of the structure of compounds occurring in nature.   As in physics, nature is called upon to corroborate assumed structures, the most important criterion in the case of proteins being furnished by their biological behavior, e.g., their reaction with enzymes.

One reason may be given for this changed attitude of the natural sciences.   The phenomena with which the scientist deals become more and more complicated.   In many cases the methods prove to be entirely too crude to aid the investigator in the

elucidation of the unquestionably very fine complexity. It is therefore this lack of adequate methods which prevents him from proceeding in a deductive manner. It is sufficient to mention the entirely different methods used in the experiments of Willstätter and collaborators for the separation of enzymes to make one realize that deductive work alone, based on older methods cannot at present help in solving our problems. This kind of work also calls for a better basic knowledge of the far reaching problem of valence which, as it appears, is extremely difficult to treat from one unitary viewpoint.

This introduction is given in order to show the difficulties which some organic chemists are facing today and to point out that these difficulties seem to be typical of science of our age. It looks as though the broadening of our knowledge of certain important natural phenomena does not proceed very quickly because the number of subordinate phenomena which can be investigated by older methods based on older conceptions is being exhausted. The axiomatic procedure in establishing a working hypothesis is on the other hand not economic since too many probabilities must be taken into consideration before one is established as a certainty by experiment. The necessity arises of inventing new standard methods that would allow the chemist to resume the deductive work. This we believe is the safest and most economic way of producing new working hypotheses. It is needless to emphasize that these remarks do not refer to organic chemistry as a whole but rather to a limited number of investigations. The latter, however, are regarded as symptomatic. It is expected that the presentation of the experimental facts in this paper will make the axiomatic tendency distinctly visible which in organic chemistry pervades also the determination of the structure of the proteins.

ALIPHATIC AND CYCLIC COMPOUNDS AMONG THE PRODUCTS OF PROTEIN DEGRADATION

Research in the composition of proteins has begun rather early. Different proteins were subjected to different treatments and the investigator attempted to obtain information on the original

material by studying its cleavage products. These early investigations, some of which possess immediate interest with view to modern protein research, helped to build firm analytical foundations. The real research into the structure of proteins starts with the establishment of the CO·NH— linkage as characteristic for proteins, pronounced nearly simultaneously by F. Hofmeister and E. Fischer. The numerous papers which were published, particularly by E. Fischer and his coworkers, prove that an immense variety of compounds was conceivable and was actually prepared on the basis of this theory. The original assumption was that the amino acids, which were regarded as the components of the proteins, are combined with each other by this CO·NH-bond in a straight chain. Thus the variety of proteins was explained by the great number of possible combinations of the various amino acids in the polypeptides and by the length of the chain of connected amino acids. These conceptions were encouraged by the fact that combinations of high molecular weight were made accessible—a very high molecular weight was generally attributed to the proteins—and that particularly higher polypeptides gave reactions resembling those of the proteins. Some of them show a characteristic behavior on salting out and give the typical color reactions. Moreover a direct experimental corroboration was obtained by the direct isolation of peptides from products of hydrolysis of proteins and by the proof that synthetically prepared peptides were split by proteolytic enzymes. The action of enzymes always was and still is regarded as the most sensitive criterion in the determination of protein structures.

It was, therefore, an interesting task to attempt the preparation of polypeptides with as many amino acids as possible in order to closely approach the tentative structure of the proteins. E. Fischer's (1) peptide with eighteen and E. Abderhalden's (2) peptide with nineteen amino acids are steps undertaken in this direction.

It is not intended to dwell upon this well known peptide theory; it is sufficient to point to its extreme usefulness. In spite of the variety of possible combinations, the idea that ultimately proteins are composed of substances of a comparatively simple

structure was agreeable since it tended to satisfy the scientific desire of simplification.

However, it was well known that the cleavage of proteins by chemical means or by peptic and tryptic enzymes has often yielded compounds which were of cyclic structure and not straight polypeptide chains. The pertaining observations were made very early and refer to the occurrence of 2,5-dioxopiperazines among the products of protein degradation. These 2,5-dioxopiperazines are composed of two amino acids; their simplest representative is glycine anhydride of the formula:

$$\text{HN} \underset{\underset{\displaystyle CH_2-CO}{}}{\overset{\overset{\displaystyle CO-CH_2}{}}{}} \text{NH}$$
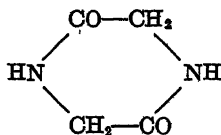
The two amino acids may be the same or different. They also may differ in their spatial configuration, thus making a great variety of these cyclic compounds possible. Among the dioxopiperazines the 3,6-diisobutyl-2,5-dioxopiperazine (leucine imide, leucine anhydride) was observed as early as 1849 by Bopp (3). Bopp digested casein with 25 per cent sulfuric acid for a day and allowed a portion of the syrup to stand for 2 months. He knew that the product obtained was a leucine derivative, but the exact formula was determined later by Erlenmeyer (4). Hlasivetz and Habermann (5) found the compound upon treatment under pressure of proteins with bromine in the presence of water, while R. Cohn obtained it on heating casein with concentrated hydrochloric acid (6).

Of course leucine imide is not the only dioxopiperazine found among the products of protein degradation. The mixed anhydrides which were found by E. Fischer and E. Abderhalden (7) were advanced as extremely important evidence in favor of the peptide theory of the proteins. Among these are the methyl dioxopiperazine (alanyl-glycine anhydride) and the glycyl-l-tyrosine anhydride from silk, glycyl-l-leucine anhydride from elastin. d-Isoleucyl-d-valine anhydride was isolated by Dakin (8) from casein. The same author obtained hydroxyprolyl-

proline anhydride from gelatin (9) which possesses an interesting structure since it contains three rings:
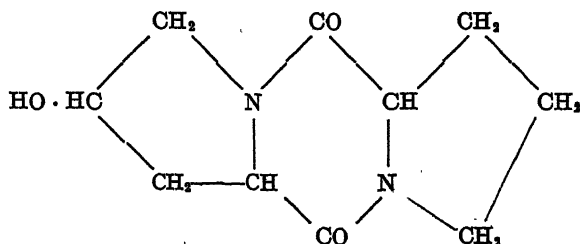
In reviewing the work that was done on the isolation of di-oxopiperazines, it is always necessary to consider the methods which were used in the individual cases for their isolation in order to be able to judge whether the particular cyclic compound was preformed in the protein or whether a secondary formation of this heterocyclic structure out of an aliphatic one is possible. In pointing to this question we are entering into the last phase of the development of the ideas of the structure of the proteins.

Although dioxopiperazines were obtained from proteins comparatively early, no systematic experiments were carried out aiming at their establishment as elementary nuclei of the protein structure. The polypeptide theory was so well supported that there was no need for the present for considering any such possibilities. It was thought that continued research along the lines of the polypeptide theory would allow to gradually elucidate the protein problem. On the other hand the possibility of the occurrence of preformed dioxopiperazines was not disregarded (10). This is shown best by a statement of E. Fischer, parts of which read in free translation as follows:

I wish to emphasize that the simple amide bond does not represent the only possible linkage within the protein molecule. On the contrary the occurrence of piperazine rings is rather probable. . . . . The numerous hydroxyl groups of hydroxyamino acids are by no means to be regarded as inert. By intramolecular formation of anhydrides they could be transformed into ether and ester groups and the variety would increase still more when one considers the polyamino acids as probable

constituents of proteins. There is no reason for a further extension of
these considerations, but I deemed it necessary to point to the various
possibilities in order to discourage too onesided ideas which would im-
pede the progress of experimental research.

NEW CONCEPTIONS OF VALENCE AND THEIR APPLICATION TO PROTEIN STRUCTURE

But it was soon realized that the assumption of polypeptide
linkages as the basic linkages of protein structure was not sufficient.
Still the degradation of proteins did not yet justify the assump-
tion of other formations. An impulse to pursue the problem in a
different direction was given by the brilliant researches of Prings-
heim, Herzog, Hess, Karrer and others in the domain of carbo-
hydrates. These investigations revealed that the colloid carbo-
hydrates are not necessarily built up of complicated original
compounds which are connected by ordinary valencies. Thus the
idea made headway that similar conditions might prevail in the
proteins. The protein chemists began to regard the proteins and
their first degradation products, the peptones and albumoses as
mixtures of a number of polypeptides. Similarly Siegfried's
kyrines could be identified by Levene and coworkers in some
instances as mixtures of polypeptides (11).

This phase of research is extremely important for the chemistry
of proteins. The assumption was made that the regular valencies
do not suffice, that we have to deal with associations, aggrega-
tions and polymerizations (12). Already in 1901 Kossel antici-
pated that the assumption of ordinary valence linkages would
not suffice, in advancing the contention that proteins are partic-
ularly inclined to formation of mixed crystals and so called solid
solutions. The forces by which the individual elementary com-
plexes are mutually attracted are of an electric nature, but we
have at present no means for their characterization. In com-
paring this valence energy with the classical one we are unable to
visualize either its capacity or its intensity factors. Therefore
new information in this realm is obtained accidentally when a
compound is prepared which shows a tendency of formation of
aggregated complexes. But with few very special exceptions we

have no means to predict what simple structure is liable to form protein-like aggregates and just what is the constitution of these associated compounds. These conceptions of aggregation are not limited to the peptides but are applied equally to the cyclic structure of the proteins.

THE THEORIES OF CYCLIC STRUCTURE OF PROTEINS

1. The pyrrole theory

The researches on polysaccharides not only helped in producing new conceptions regarding the forces of attraction between the individual primary complexes but they also fertilized the structural ideas themselves. Several theories were submitted nearly simultaneously, regarding the possible cyclic structure of proteins. N. Troensegaard regards proteins as consisting of pyrrole derivatives. The ideas that the elementary units of proteins are dioxopiperazines in addition to polypeptides is advocated particularly by E. Abderhalden. Other ring systems are also to be taken into consideration. Thus Karrer discusses the possibility of the pyrazine, imidazolone, oxazole and metoxazine oxazoline rings, while in Bergmann's theory oxazoline rings, ester peptides and recently a particular type of polymer dioxopiperazines play an important rôle. An unusual type of anhydride structure is suggested by Ssadikow and Zelinsky (14) but their suggestion is not supported by satisfactory experimental evidence.

It was known that proteins furnish pyrroles upon dry distillation and pyrrole derivatives are found among the amino acids, the primary components of the proteins (proline, hydroxyproline). Troensegaard (15) assumes, however, that the pyrrole structure is more frequent and occurs to a larger extent. He investigated the gliadin of wheat, casein, gelatin and blood proteins. In order to avoid a hydrolysis of proteins in the presence of water, the proteins were treated in non aqueous solutions. First an acetylation of the protein was carried out, then the acetylated protein was reduced with metallic sodium and amyl alcohol. With special methods a separation into several basic and acid fractions is

effected, most of which possess heterocyclic character and more specifically pyrrole structure. Troensegaard contends that the assumption of the CO·NH— linkage alone is insufficient. Neither the CO nor the NH group reacts with methyl iodide while all fractions of his protein cleavage combine with a large number of methyl groups. Thus the correctness of the polypeptide theory of the proteins must be doubted. Whenever polypeptides are found among the degradation products of proteins, it must be assumed that they are formed in a secondary manner by hydrolysis of pyrrolidone or pyrrolone (keto pyrrole) rings. According to Troensegaard the dioxopiperazine theory can be only partly correct on the following grounds. Dioxopiperazines yield oxygen free piperazines when subjected to reduction with sodium and alcohol. However, only a very small fraction of oxygen free products could be isolated by Troensegaard.

H. D. Dakin (16) also ascribes a more prominent position to the pyrrole derivatives, since comparatively large amounts of proline and hydroxyproline may be obtained from gelatin by extraction. Another interesting corroboration,—but based more on circumstantial evidence—of an extensive occurrence of pyrrole compounds in gelatin is advanced by E. Komm (17). This author finds that the reaction between tryptophane and aldehydes which may be followed up colorimetrically, is promoted by the presence of amino acids which contain pyrrole nuclei. By comparing this with the accelerating action of hydrolyzed gelatin he concludes that approximately 26 per cent of proline + hydroxyproline must be present in it which agrees very satis- factorily with Dakin's figure (14.1 per cent hydroxyproline + 9.5 per cent proline = 23.6 per cent). On the other hand the accelerating action of untreated gelatin is much more pronounced and would correspond to a content of 74 per cent of pyrrole nuclei. It is therefore suggested that labile pyrrole nuclei actually occur in gelatin but are destroyed by acid or alkaline hydrolysis. Only the comparatively resistant proline and hydroxyproline remain intact. It is important that Troensegaard's "proteols" from gelatin show a strongly accelerating influence on the tryptophane aldehyde reaction.

Although the pyrrole ring is assumed as the fundamental element of proteins, yet there is not sufficient evidence as to the composition of the individual fractions. N. Troensegaard and Eug. Fischer (18) isolated an acid from gliadin for which the following two formulas are suggested:

(1)

$$
\begin{array}{ccccc}
OC\!-\!\!-\!CCH_3 & & H_2C\!-\!\!-\!CH_2 \\
| \quad\quad \| & & | \quad\quad | \\
H_2C \quad\; C\!-\!CH_2\!-\!HC \quad\; CH\cdot COOH \\
\diagdown \;/ & & \diagdown \;/ \\
NH & & NH
\end{array}
$$

(2)

$$
\begin{array}{ccccc}
& CO & & & \\
& /\;\;\diagdown & & & \\
H_2C & \;\;C\!-\!\!-\!CH\!-\!\!-\!CH_2 \\
| & \| \quad | \quad\quad | \\
OC & \;\;C \quad CH \quad CH\cdot COOH \\
\diagdown & /\diagdown \;/\diagdown\;/ \\
& CH_2 \quad NH \quad NH
\end{array}
$$
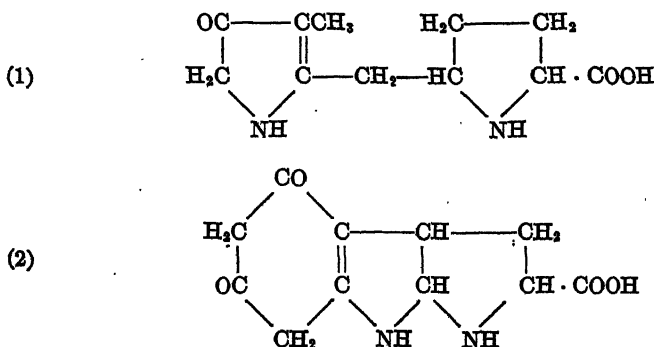
In addition biological tests were carried out and it was found that substances resulting from the reduction of acetylated proteins are more or less poisonous, showing an "alkaloid"-like behavior.

It is natural that more individual compounds of defined constitution will have to be isolated and more synthetic and enzymatic work done before Troensegaard's theory is accepted.

2. The dioxopiperazine theory

Considerable work has been done on the dioxopiperazine theory of the proteins. It was mentioned before that homolog dioxopiperazines were repeatedly isolated from degradation products of proteins. They were also the object of thorough synthetic work. It is perhaps desirable to mention some of the more important researches carried out on dioxopiperazines. Their formation from esters of amino acids was first observed by Curtius and Goebel (20). E. Fischer used this method extensively and studied the transformation of dioxopiperazines into dipeptides. The preparation of glycyl-glycine from glycine anhydride by the action of alkali is a classical example of preparation of a simple dipeptide. Numerous papers on the preparation of N, N'-diaryl-

2,5-dioxopiperazines were published by Bischoff, Widmann, Abenius and collaborators (21). With view to recent attempts (22) to isolate higher oxidation stages (tetraoxopiperazine) it may be pointed to related experiments with N,N'-diphenyl 2,5-dioxopiperazine. This compound gives 1,4-diphenyl-2-oxopiperazine when reduced with zinc dust and sulfuric acid in glacial acetic acid. The latter gives on treatment with CrO_3 in glacial acetic acid 1,4-diphenyl-2, 3-dioxopiperazine, and this is oxidizable to N,N'-diphenyltetraoxopiperazine (23). The constitution of 2,5-dioxopiperazines (24) was corroborated both by their oxidation and reduction. Thus F. Mylius (25) obtained dimethyl oxamide on oxidation of sarcosine anhydride (N, N'-dimethyl dioxopiperazine) with potassium permanganate, while E. Hoyer (26) prepared dimethyl piperazine by the reduction of alanine anhydride. Cohn (27) reduced leucine imide and observed the formation of 2,5-dibutyl piperazine. The few references represent only a part of the older literature on the subject of dioxopiperazines.

The question of primary occurrence of dioxopiperazines was raised again by E. Abderhalden and K. Funk (28). The possibility of the occurrence in silk fibroin of polymerization products of a dipeptide from alanine and glycine or the respective anhydride was discussed by R. Brill (29) on the basis of Röntgen spectrograms obtained with silk fibroin and by R. O. Herzog and M. Kobel (30). It was mentioned before that the occurrence of small elementary complexes connected by forces different from the ordinary valency was discussed by several authors. Systematic experiments aiming at the establishment of dioxopiperazine structure in proteins were undertaken by E. Abderhalden and his collaborators. Methods were devised to attack the problem simultaneously from several points. We here wish to give a logical rather than a chronological presentation, of the methods applied and the results obtained. It should be only pointed out that the first report of an attempt at the isolation of dioxopiperazines from proteins by means of additive compounds and the report of a chemical identification of a piperazine derivative obtained from a protein was given by E. Abderhalden in 1923 (31).

It was very important to obtain more information on the direct isolation of dioxopiperazines from proteins first in order to subject as large a number of proteins as possible to this investigation and secondly to secure the widest variety possible of dioxopiperazines from native proteins. As a rule the general methods used were similar in all cases. The protein was partly hydrolyzed, the product of hydrolysis concentrated in a vacuum and afterwards extracted with various solvents (ethyl acetate, methyl alcohol, acetone, chloroform, ether). The extraction with ethyl acetate gave particularly favorable results. Thus a large number of new and known dioxopiperazines was isolated.

The investigations of E. Abderhalden and coworkers show that extreme care is required in drawing conclusions from the results obtained. The utmost attention must be paid to the method of hydrolysis of the protein and it must be ascertained that the method of treatment does not include a possibility of producing anhydrides out of aliphatic peptides. In the latter case the results could be regarded as indicative of the occurrence of a certain dipeptide combination of two amino acids but would not give any evidence as to the occurrence of preformed dioxopiperazines.

It is not intended to give a list of anhydrides which were isolated by the methods described above. It should be mentioned that while in some cases the anhydrides must have been present in a preformed state, in other cases the possibility of a secondary formation cannot be disregarded. This has been shown clearly first by S. S. Grave, J. T. W. Marshall and H. W. Eckweiler (33), then by P. Brigl (34) and finally by E. Abderhalden and E. Komm (35). Thus leucyl-leucine is transformed into leucine anhydride when heated with water to 200° with a yield of more than 90 per cent of the theory. Glycyl-glycine is transformed into glycine anhydride (2,5-dioxopiperazine) when heated with 0.5 per cent hydrochloric acid under pressure at higher temperatures. Other dipeptides show a similar behavior giving anhydrides on heating with water under pressure. Some dipeptides (e.g., glycyl-tyrosine) are decomposed by this treatment. Tri- and tetrapeptides form anhydrides consisting of two amino acids

while the amino acids themselves do not give dioxopiperazines under these circumstances or only traces. Thus in judging whether dioxopiperazines are preformed in the proteins studied, we have to exclude all cases where the original method of hydrolysis might have promoted this secondary formation of dioxopiperazines from polypeptides, i.e., all investigations where heating with water or weak acids under pressure or even under reflux was used, or where an opportunity of formation of ester peptides was given, which as it is well known are readily transformed into dioxopiperazines. Treating polypeptides with 70 per cent sulfuric or with concentrated hydrochloric acid does not lead to dioxopiperazines. However, this relation between acid concentration and formation of anhydrides from peptides must be paid further attention, in consideration of the experiments of Kohler (36) who showed that leucine anhydride is obtained from leucine on heating to 220° in the presence of HCl gas and those of E. Abderhalden and K. Funk (28) who observed the formation of a small quantity of leucyl-glycine anhydride on treating leucyl glycine with 25 per cent sulphuric acid for 16 hours.

However, there is evidence that dioxopiperazines exist in a preformed state, having been extracted from proteins which were cleaved by enzymes or by concentrated acids. Salaskin (37) obtained leucine anhydride by the action of gastric juice on oxyhemoglobin and subsequent extraction with ethyl acetate. When edestin was decomposed by pancreatin and the dry residue extracted with ether glycyl proline anhydride was isolated (38):
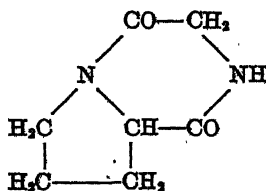
The same anhydride was obtained previously from gelatin by Levene and coworkers (30). But the value of this finding with reference to our consideration is diminished by the fact that gelatin is not a natural protein and that hydrolysis under

pressure and at a high temperature is employed in the manufacture of this product. Leucine imide was also obtained from the product of tryptic cleavage of gliadin. In addition a compound was isolated which seems to be an anhydride of leucine and glutamic acid, for which the following formulas are suggested:

```
CH₂·COOH                              COOH        CH₃
 |                                      |          /
CH₂  CO—NH                  CH₃       CO·NH·CH·CH₂·CH
 |   /        \            /           |           \
CH            CH·CH₂·CH               CH·NH          CH₂
  \          /         \              |
   NH—CO                CH₃          CH₂
                                      |
                                     CH₂
                                      |
                                     CO—
```
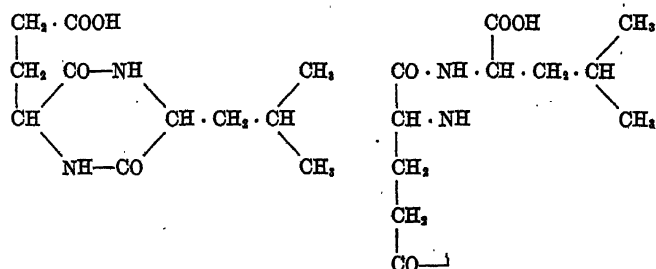
This agrees with a compound obtained synthetically by Abderbalden and Rossner.

A direct hydrolysis of casein with concentrated hydrochloric acid (without esterification) and with subsequent extraction yielded leucine anhydride (28). d-Valyl-l-leucine anhydride was obtained from casein by boiling with 5 and 10 per cent sulfuric acid and subsequent extraction of the dried product (42).

```
                      CH₃
                      /
        CO—CH·CH
       /       \    \
     HN         NH   CH₃
       \       /
        CH—CO
         |
        CH₂
         |
        CH
        / \
      CH₃  CH₃
```
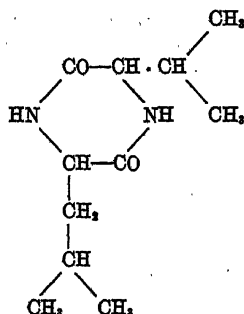
It is improbable that this procedure should have led to a secondary formation of the anhydride.

It is interesting that some dioxopiperazines are rather resistant to acids. (Their behavior toward alkalis will be dealt with later.)

E. Fischer (43) found that leucine anhydride dissolves easily in concentrated acids without decomposition. The ring is split only on prolonged heating. Thus in order to obtain leucyl-leucine the anhydride is heated in a sealed tube with hydro-bromic acid (aqueous solution saturated at 0°) at 100° for $\frac{1}{2}$ hour. Unfortunately the methods used in most of the experiments aiming at the hydrolysis of proteins, do not exclude the possibility of a secondary formation of anhydrides. They are interesting only from the standpoint of possible dipeptide combinations; actually a great variety of such new combinations was observed by E. Abderhalden and E. Komm (44) in experiments which were originally intended to adduce evidence for the occurrence of preformed dioxopiperazines in different proteins.

In discussing the isolation of dioxopiperazines from products of enzymatic cleavage, attention should be paid to the work of S. Fränkel (41a) who reports the secondary formation of anhy-drides of amino acids which are regarded as true acid anhydrides and not as dioxopiperazines.

Experiments were also carried out aiming at the isolation of dioxopiperazines by combination with certain reagents.. Al-though no definite compounds with dioxopiperazines could be isolated from hydrolyzed silk fibroin either by application of reagents which were to combine with the CO group or by those attacking the NH group (dinitrochlorobenzene) the investigations along these lines may still be successful, when a more specific reagent is found. The action of naphthalene sulfochloride on silk peptone was previously studied by E. Abderhalden and K. Funk (46).

The results obtained by Abderhalden and Stix (45) seemed to indirectly support the dioxopiperazine theory for the following reason. The NH₂ group and the NH group of dioxopiperazines are capable of reacting with dinitrochlorobenzene, while the NH- group of polypeptides is not. The authors found that the amount of dinitrochlorobenzene which entered into reaction with silk peptone was much larger than would be expected from the number of free NH₂ groups.

It is to be pointed in this connection to a paper by M.

Lüdtke (47) in which the reaction between an amino acidanhydride and phenyl isocyanate with formation of a substituted urea is described (48):

$$2\ C_6H_5NCO + HN \overset{\displaystyle CHR-CO}{\underset{\displaystyle CO-CHR}{\diagup\diagdown}} NH \rightarrow C_6H_5NHCO \cdot N \overset{\displaystyle CH_2R-CO}{\underset{\displaystyle CO-CHR}{\diagup\diagdown}} N \cdot CONH \cdot C_6H_5$$
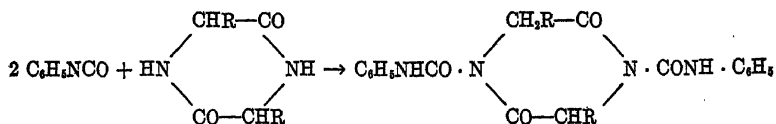
No information is available as yet as to the applicability of this reagent for the isolation of dioxopiperazines from proteins. A paper by Bergmann and Zervas (49) of recent date considers the possibility of application of aldehydes for the isolation of definite higher molecular compounds.

Although the previous experiments encouraged the conception of dioxopiperazine structure, the study of the reduction and oxidation of proteins contributed the more substantial results. Regardless of how dioxopiperazines might be combined with each other or with amino acids and polypeptides respectively, it was possible that a reduction which prevents the splitting of dioxopiperazines might produce volatile piperazines which could be driven out by steam distillation and identified in the distillate. This is actually the case and the formation of homolog piperazines from proteins is a strong point in favor of the dioxopiperazine theory. The first pertaining experiment was carried out by E. Abderhalden and W. Stix (50), who subjected silk peptone to reduction with metallic sodium and amyl alcohol. The distillate obtained gave the typical reactions of piperazines. The yield was small. But this does not indicate that dioxopiperazines are present in a small proportion since the direct treatment of dioxopiperazines in a similar manner leads to very small yields [of piperazines. This is evidently due to the low resistance of dioxopiperazines; apparently the alcohol and sodium or sodium alcoholate respectively effect cleavage to a great extent, while only a small portion is reduced.

At the same time experiments were carried out on the reduction of a number of synthetic dioxopiperazines. Thus methyl piperazine was prepared from alanyl glycine anhydride; this compound

was of particular interest since the presence of the corresponding anhydride was assumed in silk fibroin. E. Abderhalden, E. Klarmann and E. Schwab subjected glycine anhydride and leucyl glycine anhydride to reduction and obtained the respective piperazines (51). E. Abderhalden and E. Schwab (52) produced from silk peptone in addition to the methyl piperazine previously obtained the 3-methyl-6-hydroxymethyl piperazine corresponding with alanyl-serine anhydride:

$$
\begin{array}{c}
CH_3 \\
| \\
CH\!-\!CH_2 \\
\diagup \qquad \diagdown \\
HN \qquad\quad NH \\
\diagdown \qquad \diagup \\
CH_2\!-\!CH \cdot CH_2OH
\end{array}
$$
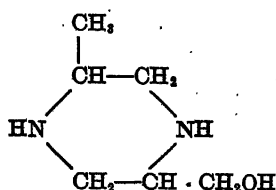
and a piperazine derived from four molecules of amino acids (two molecules of glycine, one of alanine and one of tyrosine). The latter compound was interesting since it did not give Millon's reaction in spite of the presence of tyrosine. However it was shown that the substituted piperazine resulting from the reduction of tyrosine anhydride likewise fails to give Millon's reaction. This points to the fact that the presence of the piperazine nucleus interferes with this test.

The same authors reduced directly untreated gelatin with sodium and alcohol (ethyl or amyl) and by using phenyl isocyanate for the separation of the individual fractions isolated piperazines which are derived from hydroxyproline glycine anhydride (I) and proline leucine anhydride (II) respectively:

$$
\begin{array}{cc}
\begin{array}{c}
CH_2\!-\!CH_2 \\
\diagup \qquad\quad \diagdown \\
HN \qquad\qquad N\!-\!CH_2 \\
\diagdown \qquad\quad \diagup \quad\ | \\
CH_2\!-\!CH \qquad\ | \\
| \qquad\qquad \\
CH_2\!-\!\!-\!CHOH
\end{array}
&
\begin{array}{c}
CH_3 \\
\diagdown \\
\quad CH \cdot CH_2 \cdot CH\!-\!CH_2 \\
CH_3 \diagup \qquad\qquad\qquad \diagdown \\
\qquad HN \qquad\qquad\quad N\!-\!CH_2 \\
\qquad\ \diagdown \qquad\quad \diagup \\
\qquad\quad CH_2\!-\!CH \\
\qquad\qquad\qquad CH_2\!-\!CH_2
\end{array} \\
I. & II
\end{array}
$$
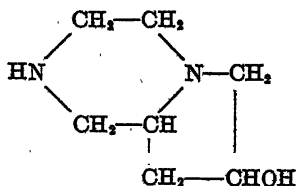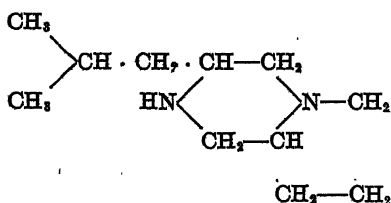
This experiment is however not convincing, since a secondary

formation of dioxopiperazines in gelatin is possible with view to the methods used in its preparation (53).

The behavior of amino acids and polypeptides in the presence of reducing agents of the type described was likewise studied and it was found that in no case the formation of a piperazine takes place. On reduction of dipeptides the formation of amino alcohols with partial deamination is observed. Thus the reduction of leucyl-glycine leads to γ-methyl-α-hydroxymethyl butyl-amine (leucinol):

$$\begin{array}{c} CH_3 \\ {>}CH \cdot CH_2 \cdot \underset{\underset{NH_2}{|}}{CH} \cdot CH_2OH \\ CH_3 \end{array}$$
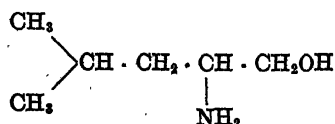
At the same time caproic acid forms (54). Upon reduction of d,l-alanyl-glycine α-hydroxymethylethylamine (I) and propionic acid were obtained, while glycyl-glycine yielded hydroxyethyl-amine (II).

$$\underset{\underset{NH_2}{|}}{CH_3 \cdot CH} \cdot CH_2OH \qquad \underset{\underset{NH_2}{|}}{CH_2} \cdot CH_2OH$$

$$(I) \qquad\qquad (II)$$
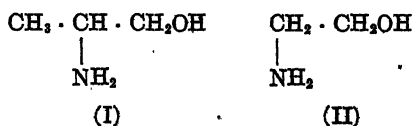
E. Abderhalden and coworkers also took the possibility of primary occurrence of piperazine nuclei into consideration and synthesized a considerable number of piperazine-amino-acid compounds. These experiments will be dealt with later in connection with the other studies on model compounds.

The formation of piperazine from proteins upon reduction certainly points to the probability of the primary occurrence of dioxopiperazines. Abderhalden and coworkers carried out another series of experiments aiming at the establishment of their presence in proteins by oxidation methods. Their results were corroborated by those of S. Goldschmidt and Ch. Steigerwald.

The use of the oxidation method is also based upon a different behavior of peptides, amino acids and dioxopiperazines to oxidizing agents. Oxidations of proteins with various oxidizing

agents were carried out in the past. The results of Loew (35) and Kutscher and Schenk (56) are particularly interesting. These authors found oxamide and oxaminic acid among the products of oxidation of proteins with potassium permanganate. F. Mylius' oxidation of sarcosine anhydride was mentioned before. In order to ascertain whether these oxidation products were indicative of a particular structure E. Abderhalden with E. Klarmann and E. Komm (57) subjected dipeptides, their corresponding anhydrides and silk peptone to oxidation. Zinc permanganate was used which is more convenient than the potassium salt.

All dioxopiperazines yielded oxamide. The dipeptides were decomposed with the exception of glycyl-glycine which also yielded oxamide. In the case of the simplest dioxopiperazine the oxidation takes place according to the equation

$$
\begin{array}{cc}
\text{NH} & \text{NH}_2 \\
\diagup \quad \diagdown & \diagup \\
\text{OC} \qquad \text{CH}_2 & \text{CO} \\
\mid \qquad \mid \quad + 5\ \text{O} \rightarrow \mid \qquad + 2\ \text{CO}_2 + \text{H}_2\text{O} \\
\text{H}_2\text{C} \qquad \text{CO} & \text{CO} \\
\diagdown \quad \diagup & \diagdown \\
\text{NH} & \text{NH}_2
\end{array}
$$
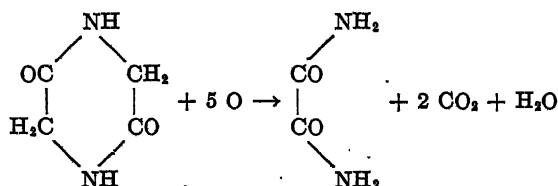
Similarly the oxidation of glycyl-glycine is represented by the equation:

$$
\begin{array}{cc}
\text{NH}_2 & \text{NH}_2 \\
\mid & \mid \\
\text{CH}_2 \cdot \text{CO} \cdot \text{NH} \cdot \text{CH}_2 \cdot \text{COOH} \quad + 5\text{O} \rightarrow & \text{CO} \cdot \text{CO} \cdot \text{NH}_2
\end{array}
$$

It was shown later by E. Abderhalden and E. Komm (58) that oxamide is obtained also on oxidation of polypeptides which contain the glycyl-glycine group. All the other polypeptides did not give oxamide on oxidation, while all amino acid anhydrides (dioxopiperazines) and their $0,0'-$ or $N,N'-$ substituted derivatives yielded oxamide in a satisfactory quantity.

This method was also used in the oxidation of proteins. Some proteins are very difficultly attacked by permanganate. But oxamide was obtained from gelatin, blood globulin, egg albumen and caseinogen (Hammarsten) in addition to the silk peptone previously mentioned.

It may be added in this connection that oxamide results also on oxidation with hydrogen peroxide.

The experiments of S. Goldschmidt and Ch. Steigerwald (59) belong in this group. The authors realized that only mild methods should be used in the experiments aiming at the elucidation of the structure of the protein molecule. They found that proteins are attacked by alkali hypobromite at $0°$; different proteins react with varying amounts of hypobromite during the same length of time. Thus a series of titration curves may be obtained, each being characteristic of a given protein. The age of a protein solution is discernible by means of the titration curve, freshly prepared solutions of some proteins absorbing more hypobromite than older ones. It is assumed that the $NH-$ group is attacked by hypobromite according to the equation:

$$-CO \cdot NH \cdot CH\big\langle \rightarrow CO \cdot NBrCH\big\langle \xrightarrow{NaOH} CON:C\big\langle \xrightarrow{H_2O} CONH_2 + OC\big\langle$$
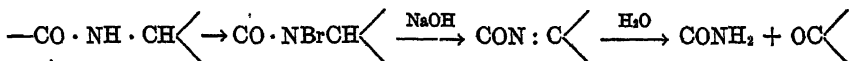
On the other hand it was found that it is not the $NH-$ group of peptides which reacts in this manner, since glycyl-glycine is completely split by hypobromite giving ammonia and oxalic acid. The probable course of this reaction is the following one:
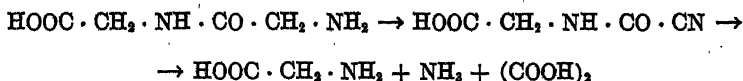
$$HOOC \cdot CH_2 \cdot NH \cdot CO \cdot CH_2 \cdot NH_2 \rightarrow HOOC \cdot CH_2 \cdot NH \cdot CO \cdot CN \rightarrow$$

$$\rightarrow HOOC \cdot CH_2 \cdot NH_2 + NH_3 + (COOH)_2$$

However, dioxopiperazines react with hypobromite. From glycine anhydride 4-imidazolone-2-carboxylic acid is obtained in a 20 per cent yield; it is extremely sensitive toward alkalis.
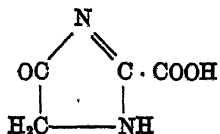
Alanine anhydride shows the same behavior. It is concluded from the analogy of behavior of dioxopiperazines and the proteins, that it is very probably the dioxopiperazine ring in the proteins that reacts with hypobromite.

Finally it should be pointed to a paper by Z. Stary (60) who studied the action of bromine in glacial acetic acid and of hydrogen peroxide in a 4N solution of sulfuric acid on human hair and the enzyme action (trypsin) on the resulting degradation products. He mentions that the primary existence of dioxopiperazine nuclei is to be taken into consideration which subsequently under the influence of splitting agents form polypeptides.

Thus the oxidation experiments also suggest the primary occurrence of dioxopiperazines in proteins. It should, however, be kept in mind that oxamide, the formation of which is the criterion of most of the oxidation experiments is also produced on oxidation of some polypeptides, although of a highly specific class. It would be interesting to carry out oxidations of proteins which contain only little glycine in order to exclude the possibility of the interference by the presence of glycyl glycine. The establishing of a quantity of oxamide larger than would correspond to the amount of glycine present in the protein, would show more convincingly that preformed dioxopiperazines are present. However, it must not be disregarded that comparatively few oxidation experiments were carried out on heterocyclic compounds and that oxamide might possibly be formed by oxidation of ring compounds different from dioxopiperazines.

Another series of experiments aiming at the comparative investigation of dioxopiperazines, peptides, amino acids and proteins or their cleavage products respectively, includes the application of color reactions which would be specific for one class but would not be given by other classes. It is perhaps useful to mention in this connection that group distinction in protein chemistry by means of color reactions is well known and extensively applied. Thus ninhydrine (triketohydrindene hydrate) gives color reactions on heating with amino acids, peptones and soluble proteins. On the other hand the biuret reaction is given by proteins, some higher peptides and peptones but not by amino acids. (It is not important that it is given by amino acid amides, since one does not take their occurrence in proteins into consideration.) Both reagents do not give any reaction with

dioxopiperazines. On the other hand precipitated moist copper oxide reacts readily with amino acids, and peptides but not with dioxopiperazines. The combination of the three tests is utilized for a qualitative analysis of a protein compound. But it is not suited for the determination of the occurrence of a dioxopiperazine in a protein.

E. Abderhalden and coworkers started a search for reagents which would be more or less specific for dioxopiperazines. Particularly aromatic nitro compounds (carbonyl reagents) were found to give characteristic reactions with dioxopiperazines. In order to carry out the reaction the reagent and sodium carbonate (in one case sodium alcoholate) is added to the tested solution and boiled for a short time. The following reagents were suggested out of a number of polynitro derivatives: picric acid, m-dinitrobenzene, 1,3,5-dinitrobenzoic acid, m-dinitrostilbene.

Picric acid was used originally by Jaffe as reagent for creatinine (61). T. Sasaki showed that amino acid anhydrides which contain glycine as a constituent also give a positive reaction. However, a number of different compounds, e.g., hydantoins, glucose, malonic ester, acetic ester also give a positive reaction. H. A. Dox (63) showed that the same reaction is given by barbituric acid, but it is negative with mono- or dialkyl barbituric acids. E. Abderhalden realized that the reaction is not absolutely specific, but most of the above named cases could be excluded, since the occurrence of these compounds in proteins is highly improbable. He therefore investigated systematically in collaboration with E. Komm and E. Schwab (64) a considerable number of amino acid anhydrides and found that all with the exception of l-leucyl-d-leucine anhydride gave a positive reaction with picric acid and sodium carbonate. No amino acid, polypeptide or biogenic amine, but all peptones and the majority of proteins gave a positive reaction. It is contended that the negative reaction with some of the proteins is due to their insolubility, although some of them contain reacting groups as shown by their behaviour toward m-dinitrobenzene, the reagent of v. Bittó.

Similar results were obtained with m-dinitrobenzene. All amino acid anhydrides gave a positive reaction, similarly all peptones and proteins, e.g., silk peptone, plant casein, blood globulin, keratin, ricin, legumin, vitellin, gelatin. Neither amino acids nor peptides give the reaction. l-Leucyl-d-leucine anhydride also gives a negative reaction. The authors prefer this reagent to picric acid, since a much more distinct change of color is observed and the reaction takes place more readily.

It is important to note that compounds which are regarded as N,N'— substituted dioxopiperazines give positive color reactions, while the reactions of the O,O'— substituted derivatives are negative.

The results obtained with 1,3,5-dinitrobenzoic acid are largely the same. However, some irregularities were observed. Thus it was found that cystine, cysteine, diglycyl-cystine, diglycyl-piperazine and dialanyl-piperazine gave positive reactions with m-dinitrobenzoic acid although they do not contain the dioxopiperazine ring. m-Dinitrostilbene (2,4) behaves like the other reagents except that no reaction is obtained with cysteine, cystine or its derivatives. On the other hand a few proteins which gave positive reaction with the other reagents give a negative reaction here, e.g., plant casein or egg albumen.

These color reactions represent in spite of their not entirely specific nature another support of the dioxopiperazine theory. Although for themselves they would not be absolutely conclusive yet in conjunction with the direct isolation of anhydrides, and the products isolated from the oxidation and reduction of proteins they contribute to the strengthening of the dioxopiperazine theory.

It must not be disregarded, however, that the color tests when carried out in a modified manner according to Brand and Sandberg (65) fail to give positive results with proteins, while a positive reaction is obtained with ordinary carbonyl compounds and dioxopiperazines using the same method. This means possibly that the short heating with sodium carbonate solution as required in Abderhalden's specifications of the method first

loosens the combination of the dioxopiperazines with some other compounds before the free dioxopiperazine is able to give the positive color reaction.

It is obvious that the investigator attempts to synthesize the compounds which he determines in the natural products. This is particularly important here. It is true that there is considerable evidence for the existence of preformed dioxopiperazines in proteins. But after all they represent only small complexes. Where is the connection with the compounds of high molecular weight?

It was mentioned before that, on the basis of the polypeptide theory, attempts were made to arrive at compounds of high molecular weight similar to proteins by connecting individual amino acids to long chains. Similarly, Abderhalden attempted first to prepare compounds consisting of dioxopiperazines and amino acids, in order to realize a possibility of connecting a number of dioxopiperazines with each other. While E. Fischer's method of building long polypeptide chains is comparatively simple, no such method could be devised here. It is necessary that much more information be gathered on the chemical properties of dioxopiperazines, before this problem can be successfully handled. On the other hand it does not further the cause of science when hypotheses are advanced involving forces of attraction which we cannot clearly deal with. Such hypotheses are not justified unless all means of experimental approach have been exhausted and the investigator realizes that he is confronted with entirely unknown conditions. Although it is impossible at the present time to make any statements with reference to the problem of valence within the protein molecule on the basis of analytical research, one may say safely that there is still a considerable amount of experimental work to be done and there is no reason except the still disputable x-ray evidence why a careful study of compounds with true chemical linkages should not be carried out.

Of course such synthetic compounds are tentative models. There is experimental evidence for the existence of amino acid

anhydrides containing more than two amino acids, but the
structural formula attributed to them is a probable and not an
established one. Some of these compounds which were isolated
from native proteins and their tentative formulas shall be given
as examples: a compound consisting of three molecules of l-proline
(66) with the tentative formula:
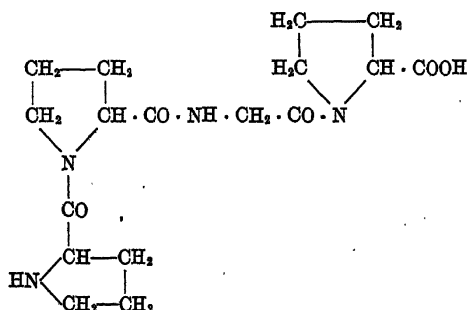
an anhydride consisting of two molecules of l-proline, one mole-
cule of hydroxyproline and one molecule of glycine:
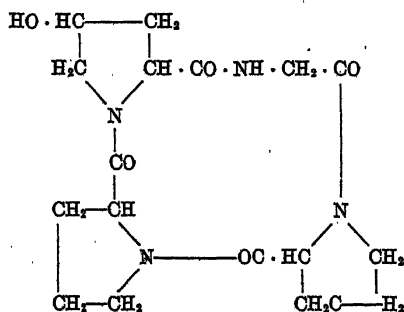
Both were obtained from keratin (goose feathers). From casein
an anhydride was produced consisting of one molecule of
d-alanine, one molecule of l-leucine and one of l-proline (67) for
which the following formulas are suggested:

$$CH_3 \qquad CH_2\!\!-\!\!CH_2 \quad CH_2\!\!-\!\!CH_2 \qquad\qquad CH_3$$

$$CH\!\!-\!\!CO \qquad\qquad CH_2 \quad CH_2 \cdot CO \cdot NH \cdot CH \cdot CH_2 \cdot CH$$

$$HN \qquad N\!\!-\!\!OC \cdot CH \quad CH_2 \qquad N \qquad\qquad CH_3$$

$$CO\!\!-\!\!CH \qquad NH \qquad\qquad CO$$

$$CH_2 \qquad\qquad CH \cdot NH\!- \qquad -CO$$

$$CH \qquad\qquad\qquad CH_3$$

$$CH_3 \qquad CH_3$$
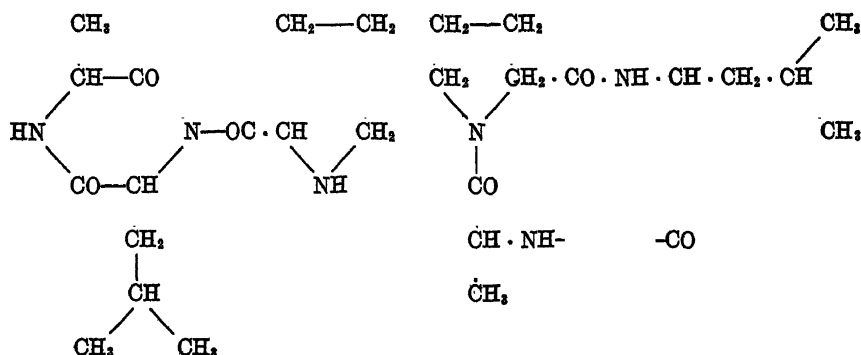
It was mentioned before that the presence in silk fibroin of an anhydride containing two molecules of glycine, one molecule of d-alanine and one of l-tyrosine is assumed by Abderhalden and Schwab (52).

$$CH_3$$

$$NH_2 \qquad CH\!\!-\!\!CO$$

$$HO \cdot C_6H_4 \cdot CH_4 \cdot CH \cdot CO \cdot N \cdot N \qquad N \cdot OC \cdot CH_2 \cdot NH_2$$

$$CO \cdot CH_2$$
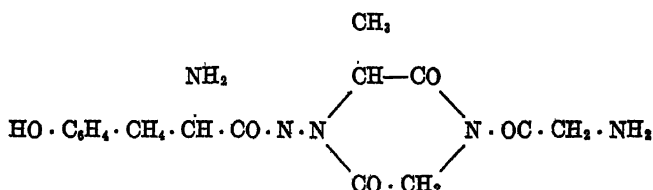
The number of compounds of this type is much larger (68) but the few examples cited, suffice to produce an idea as to the assumed nature of combinations of dioxopiperazines and amino acids.

It is noteworthy that upon addition of alkali the amount of NH_2 — nitrogen increases, thus indicating that the ring is opened on one side by this treatment.

All these compounds were found in native proteins. On the other hand the problem was approached also synthetically. The first methods used resembled those of E. Fischer for the preparation of polypeptides. The direct condensation with halogen acyl halides and subsequent amination did not lead to the desired products. E. Abderhalden and E. Klarmann found that when water is excluded a condensation may take place at a higher temperature with formation of the corresponding di-

(halogen acyl)-dioxopiperazines. Thus chloroacetyl chloride reacts at 160° with glycine anhydride in nitrobenzene solution:

$$Cl \cdot CH_2 CO \cdot N \underset{CH_2-CO}{\overset{CO-CH_2}{\diamond}} N \cdot CO \cdot CH_2Cl$$
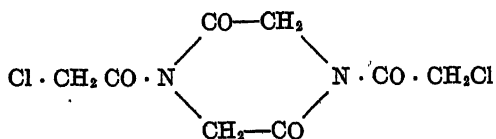
Similarly di-(α-bromoisocaproyl)-dioxopiperazine is obtained (69) on boiling glycine anhydride with the acid chloride. However, it was impossible to replace the halogen by the NH_2 group since in the presence of NH_3 cleavage takes place in all solvents with formation of the acid amide and free dioxopiperazine. It is perhaps interesting to add that according to Karrer (70) these two compounds might be O,O'— substituted derivatives. The easy formation of acid amides from an ester-like compound would then be intelligible since it is known that esters readily produce acid amides in contact with ammonia. On the other hand compounds were prepared synthetically (see page 84) which probably have the structure

$$\overset{-CH_2}{\underset{-CO}{\diagup}} N-CO-$$
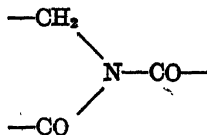
and are rather resistant to ammonia.

The synthetic preparation of such compounds is important with view to the reaction toward enzymes. The study of the behavior of dioxopiperazines toward enzymes shows so far, that no cleavage takes place. These investigations will be dealt with later. But it is possible that the nature of the dioxopiperazine nucleus might be so changed by the introduction of an amino acid rest that the resulting compound would be attacked by enzymes. It was therefore attempted to prepare compounds of amino acids and dioxopiperazines using different methods. Before we turn to these investigations we shall first consider the nature of the dioxopiperazine nucleus itself.

The structural formula of the dioxopiperazine nucleus permits the following tautomer structures to be assumed (71):

```
       R                 OH  R              OH  R
       |                  |  |               |  |
   CO—CH               C——CH              C══C
  /      \            //     \           /      \
HN        NH         N        N        HN        NH
  \      /            \      //          \      /
   CH—CO               CH——C              C══C
       |                |   |               |  |
       R                R  -OH             R   OH
       I.                  II.                III.
```
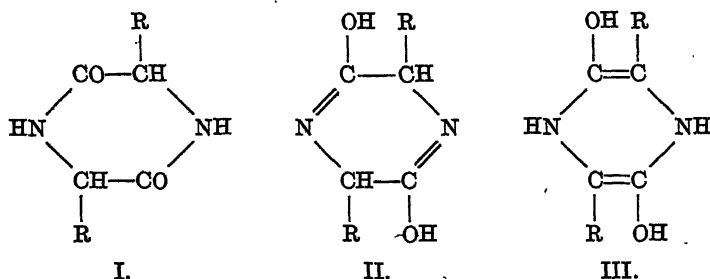
Actually compounds were obtained under certain conditions which possess an outspoken unsaturated character (72). Thus when glycine anhydride is heated with glycerol in the presence of tyrosine to 190–200°C. for five hours a compound is isolated which has the empirical composition of glycine anhydride. However, in contrast to this, it immediately decolorizes permanganate, gives a positive xanthoprotein reaction and readily allows the introduction of methyl groups by means of diazo methane. The authors assume the formula III as the most probable for this compound. The same treatment in the absence of tyrosine leads to a compound which originally possesses this unsaturated nature but loses it in the course of purification. Similarly heating of d,l-leucyl-glycine anhydride with glycerol in the presence of tyrosine leads to a substance which shows all properties of an unsaturated compound (73). The rôle of tyrosine in this connection is not entirely clear. Rearrangement of the enol form of 2,5-dioxopiperazine into its keto form takes place by heating in aqueous solution to 90–100° (74).

Another method of obtaining dioxopiperazines in the enol form is heating of the respective dipeptides with diphenylamine (75). The enolic anhydrides of d,l-leucyl-glycine, d,l-leucyl-d,l-valine, d,l-alanyl-d,l-alanine, d,l-leucyl-d,l-leucine were prepared in excellent yields. Glycyl-glycine and glycyl-alanine showed a behavior different from that of the other dipeptides. The first gives a difficultly soluble, probably polymer compound with the empirical composition of glycine anhydride. While dioxo-

piperazines in keto form cannot be transformed into the enol form by heating with glycerol or diphenylamine, this transformation is effected by heating with aniline (79).

An interesting proof for the assumed structure (I):

$$
\begin{array}{cc}
\begin{array}{c}
\mathrm{OH\ R} \\
|\quad| \\
\mathrm{C}=\mathrm{C} \\
\diagup\qquad\diagdown \\
\mathrm{HN}\qquad\mathrm{NH} \\
\diagdown\qquad\diagup \\
\mathrm{C}=\mathrm{C} \\
|\quad| \\
\mathrm{R\ OH}
\end{array}
&
\begin{array}{c}
\mathrm{CH_3\quad CH_3} \\
\diagdown\quad\diagup \\
\mathrm{OC}-\mathrm{C} \\
\diagup\qquad\diagdown \\
\mathrm{HN}\qquad\mathrm{NH} \\
\diagdown\qquad\diagup \\
\mathrm{C}-\mathrm{CO} \\
\diagup\qquad\diagdown \\
\mathrm{CH_3\quad CH_3}
\end{array} \\
\text{(I)} & \text{(II)}
\end{array}
$$
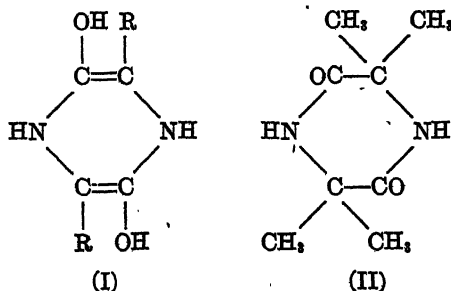
was given by the preparation of the anhydride from α-amino-isobutyryl-α-aminoisobutyric acid (II). This anhydride cannot exist in the previously mentioned enol form on account of its particular structure. Actually the synthetized product gave all anhydride reactions, but unlike the other anhydrides prepared by the same method, it failed to give the xanthoprotein reaction and did not decolorize permanganate. Since it was possible to obtain sarcosine anhydride in the unsaturated form (79), the existence of the $-\mathrm{C}=\mathrm{C}-$ linkage in enolic dioxopiperazines seems to be established.

Methods were devised which also allow to distinguish between these tautomer compounds in a physical way. E. Abderhalden and R. Haas found that a number of amino acid anhydrides give a characteristic absorption in the ultraviolet, the enol form showing a more pronounced absorption than the keto form. It is important that some proteins also absorb in the ultraviolet. The interesting fact was found that the absorption spectrum of an amino acid depends on the method of its preparation. Thus alanine precipitated from its aqueous solution with alcohol behaves differently from alanine obtained by direct crystallization. Possibly this observation is related to that of E. Fischer (77) that the precipitation with alcohol influences the amino

acid so that a preparation of the acid chloride is made possible, while no acid chloride can be made from an acid obtained directly from water. In one case, the enol-keto rearrangement could be observed spectroscopically in d,l-leucyl-glycine anhydride and it was found that the rearrangement is complete after 8 hours (78).

These results are interesting from two points of view. First the ultraviolet absorption of dioxopiperazines may be adduced as evidence for their occurrence in proteins. Secondly it was possible to show that there are labile enol modifications which show a measurable rearrangement.

However, it must be taken into consideration that according to W. Stenström and M. Reinhard (80) the ultraviolet absorption of the proteins is due to the aromatic amino acids. Y. Shibata and T. Asahina (81) also investigated the question of desmotropy of dioxopiperazines. Their investigations are based on those of J. N. Hartley (82). This author studied the possible desmotropy of cyanuric acid for which the following formulas were possible:

I.   II.
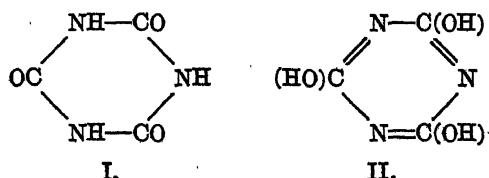
The second formula resembling to a certain degree the benzene nucleus, it was expected that the compound would show absorption in the ultraviolet. Since this was not the case, the conclusion was drawn that the enol tautomer does not exist. Shibata and Asahina prepared dioxopiperazines (glycine anhydride, alanine anhydride, sarcosine anhydride) by heating the corresponding amino acids in glycerol according to Balbiano and Trasciati (83) and Maillard (84) and none of the compounds studied showed an absorption spectrum; the conclusion was therefore drawn that they all exist in the keto form only. There seems to be a discrepancy between these observations and those of Abderhalden and Haas. In any case the continued research

into the desmotropy of dioxopiperazines with application of optical methods seems to promise interesting results.

Some interesting observations have been made on the dioxo-piperazines which later will be possibly utilized for structural investigations. Thus the refractive index of the enol form was found to be higher than that of the keto form (85). The optical rotation of solutions of dioxopiperazines decreases under the influence of Röntgen rays and ultraviolet rays, while that of the corresponding peptides remains unchanged. Probably an oxidation takes place due to the formation of ozone under the influence of irradiation.

It is very important that the cleavage of enolic dioxopiperazines leads to unsaturated dipeptides (86). Glycyl-glycine prepared from enolic glycine anhydride behaves differently from that obtained from the keto form. It decolorizes permanganate in the cold, gives a positive xanthoprotein reaction, a positive ninhydrine and negative anhydride reaction and dissolves in aqueous sodium hydroxide giving intensely yellow colored solutions. The yellow color disappears on heating. A similar behavior is shown by the unsaturated form of d,l-leucyl-glycine, which may be prepared by splitting the corresponding anhydride or by heating ordinary leucyl-glycyl-leucine with diphenylamine to 200°. The two forms of leucyl-glycine show a different behavior toward alkali. While the saturated form is entirely unchanged in the presence of 0.2 N sodium hydroxide after a lapse of 15 hours, the unsaturated form is split quantitatively after 9 hours.

The observations made on the introduction of the acyl rest into benzylated glycine anhydride would also indicate that the enolic form of dioxopiperazines possesses a higher reactivity. While it was previously found that glycine anhydride must be heated with chloroacetyl chloride in the presence of nitrobenzene to 160° in order to effect a substitution, the dichloroacetyl compound results on heating the dibenzyl dioxopiperazine with chloroacetyl chloride at water bath temperature. According to Karrer (70) the dibenzyl compound is derived from an enolic form for which the following formula is suggested:

$$O \cdot CH_2 \cdot C_6H_5$$

$$
\begin{array}{c}
C \text{---} CH_2 \\
\end{array}
$$

$$
\begin{array}{ccc}
 & C \text{---} CH_2 & \\
N & & N \\
 & CH_2 \text{---} C & \\
\end{array}
$$

$$O \cdot CH_2 \cdot C_6H_5$$
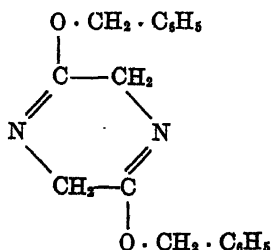
The double bonds are not in the position assumed by Abderhalden and Schwab. With view to these experiments it is not excluded that heating 2,5-dioxopiperazine in nitrobenzene likewise leads to an enol structure, possibly different from that which results when dipeptides are treated in diphenylamine, or dioxopiperazines in aniline.
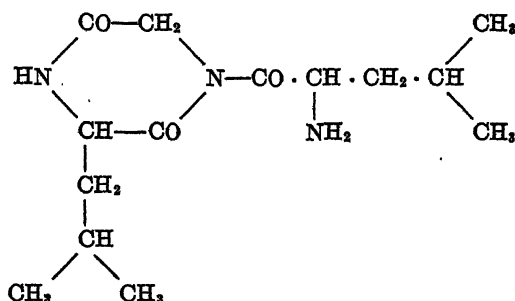
The existence of labile polypeptides with enolic structure opens a new field of possible relations to amino alcohols, amino aldehydes, etc., and generally to the reactions of proteins in metabolism. It is not excluded that the view of the total degradation of proteins introduced into metabolism will have to be changed. Similarly it might be found that compounds other than amino acids are being absorbed in the gastro-intestinal tract. It is held now with view to the experiments with ferments that the stable modification of dioxopiperazines occurs in proteins which are resistant to the action of enzymes, while the labile enolic modification occurs in the proteins accessible to enzymatic cleavage. It is also possible that the denaturation of proteins by heat depends upon the transition from the enol into the keto modification (32).

The researches which were carried out by L. Balbiano and later by L. C. Maillard and created new possibilities for the synthetic preparation of dioxopiperazines and their derivatives deserve mention in this place. Every such contribution to the synthesis is highly welcome to the investigator of proteins, since it opens urgently needed new sources of supply, the older methods of preparation of dioxopiperazines being none too convenient. Balbiano heated glycine in a sealed tube to 150–170° and ob-
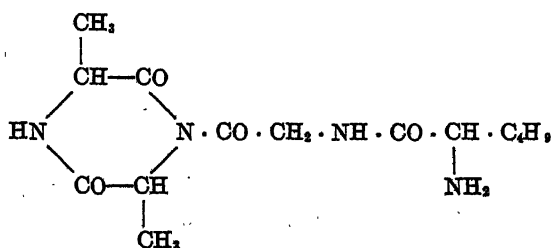
tained a keratin-like compound which he regards as a polymer glycine anhydride $(C_2H_3ON)_x$. It is interesting that on heating glycyl-glycine with diphenylamine to 185–190° an extremely difficulty soluble compound results (88), which likewise is regarded as a polymer of glycine anhydride. Maillard studied the influence of hot glycerol on amino acids. According to his investigations the composition of the products depends upon the amount of glycerol. Thus either glycine anhydride or the tetrapep- tide triglycyl-glycine may form. This formation of a tetrapeptide may have some bearing upon Bergmann's researches which will be dealt with later. In addition a number of different glycine combinations is obtained, e.g., the tripeptide diglycyl-glycine and the hexapeptide pentaglycyl-glycine. Likewise the formation of a polymer compound, possibly cycloheptaglycyl-glycine $(C_2H_3ON)_8$ is observed. This compound seems to · possess remarkable properties. It dissolves in concentrated acids and gives upon dilution solutions which first stay clear and give a positive biuret reaction; on standing the solution becomes turbid and the biuret reaction becomes negative. Kaito Shibata (89) also reports the formation of dioxopiperazines on heating proteins with glycerol.

This method was used by Abderhalden and collaborators in the preparation of simple and mixed anhydrides (e.g., d,l-leucyl-glycine anhydride from leucine and glycine) (90). Heating of dipeptides with glycerol likewise leads to anhydrides (91). But it is more important that compounds were obtained synthetically which in all probability consist of combinations of anhydrides and amino acids and resemble those obtained by the degradation of proteins. Thus on heating of d,l-leucine with d,l-leucyl-glycine anhydride the leucyl-(glycyl-leucine) or leucyl-(leucyl-glycine) anhydride respectively is obtained. Similarly heating the tripeptide leucyl-glycyl-leucine leads to a compound to which the following constitution or its tautomer is attributed:

$$\begin{array}{c}
\underset{\underset{\displaystyle\underset{\displaystyle\underset{\displaystyle CH_2 \quad CH_3}{\diagup \diagdown}}{CH}}{\underset{\displaystyle CH_2}{|}}}{HN\diagdown \underset{\displaystyle CH-CO}{\diagdown}} \overset{\displaystyle CO-CH_2}{\diagup} N-CO\cdot CH\cdot CH_2\cdot \overset{\displaystyle CH_3}{\underset{\displaystyle CH_3}{CH}} \\
NH_2
\end{array}$$

The same compound may be obtained from the methyl ester of leucyl-glycyl-leucine (91). The positive ninhydrine and anhydride reactions confirm its structure. The biuret reaction is, negative. Also a dipeptide was brought into reaction with an amino acid anhydride. On heating alanine anhydride directly with leucyl-glycine in aniline to 200°, a compound with the probable structure:

$$HN\diagdown \overset{\displaystyle \overset{\displaystyle CH_3}{\underset{\displaystyle CH-CO}{|}}}{\diagup}\diagdown N\cdot CO\cdot CH_2\cdot NH\cdot CO\cdot \underset{\displaystyle NH_2}{\overset{\displaystyle CH}{\underset{\displaystyle |}{CH}}}\cdot C_4H_9$$

was obtained. This compound also gives positive ninhydrine and anhydride reactions and no biuret reaction. In aqueous solution it behaves like a colloid since it first dissolves clearly and then flocculates on standing. It is intended to study the action of enzymes on compounds of this type.

It was mentioned previously that in addition to the dioxopiperazine ring, the occurrence of the piperazine ring itself was taken into consideration. Numerous experiments were carried out on the introduction of amino acids and polypeptide rests into the piperazine ring and its alkylated derivatives. The compounds thus obtained were exposed to the action of enzymes (92) but in no case a cleavage was observed. The formula of diglycyl-

leucyl-2,5-dimethyl piperazine prepared by Abderhalden and Kohl-Egger (93) is given here in order to convey a general idea of the structure of compounds of this group:

$$
\begin{array}{c}
\text{CH}_3 \\
|
\end{array}
$$
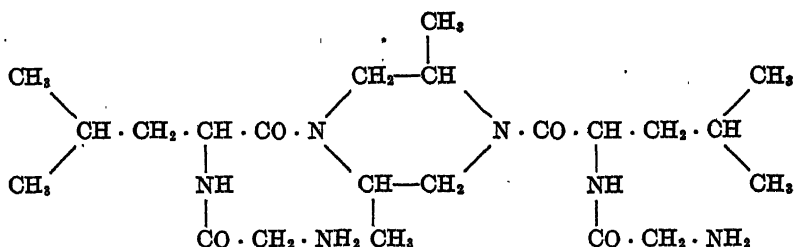
This compound is not attacked by yeast juice.

Abderhalden and coworkers attempted to find more data for the distinction of peptides and anhydrides in proteins in addition to the tests already mentioned (94). The adsorption experiments are very interesting. Solutions (0.05 N) of dipeptides and dioxopiperazines were shaken with animal charcoal for 15 minutes; 8.45 per cent of glycyl-glycine or 7.37 per cent of alanyl-glycine are removed by this treatment in contrast to as much as 38.93 per cent glycine anhydride and 37.66 per cent alanyl-glycine anhydride. Further it was found that precipitating agents act differently on amino acids, polypeptides and anhydrides. The comparative cleavage of peptones and proteins by alkali also gives noteworthy results. Experiments on the action of alkali and acid on dioxopiperazines and dipeptides with control of pH were conducted first by M. Lüdtke (95). Abderhalden and Haas showed (96) that the cleavage of dioxopiperazine is an equilibrium reaction:

$$\text{Dioxopiperazine} + \text{H}_2\text{O} \rightleftarrows \text{dipeptide}$$
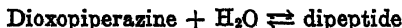
Accordingly, it is possible to effect a shift either by addition of freshly precipitated Cu(OH)$_2$ which is bound by the free peptide, or by treatment with yeast juice which produces enzymatic cleavage of the peptide. While glycyl-glycine and leucyl-glycine are perfectly stable at pH 12.4, the corresponding dioxopiperazines are split at this and lower pH. The action of alkali on

silk peptone is accompanied by a slow decrease of pH, but simultaneously the strong picric acid and dinitrobenzoic acid reaction is considerably weakened. The same phenomenon is observed on casein Hammarsten, casein Osborne and beef blood serum. This points to the splitting of preformed dioxopiperazine rings while the alkali is neutralized by the carboxylic group of the peptide formed. There is also a close relation between the picric acid reaction, the formation of free NH_2 groups and the yield of oxamide (97).

It is perhaps desirable to mention in this connection the experiments of Fischer and Schrauth. These authors found that the ease of cleavage of a dioxopiperazine by alkali depends upon the amino acids which constitute the particular dioxopiperazine. While anhydrides with glycine are easily split, neither leucine anhydride nor valine anhydride are attacked by dilute alkali even on standing for 10 days at 37°. In accordance with these findings tyrosine anhydride is split much more difficultly than glycyl tyrosine anhydride. With view to the objections to the dioxopiperazine theory which will be dealt with later it might be interesting to subject one of the proteins from which leucine anhydride was isolated by enzymatic or acid hydrolysis (e.g., hemoglobin) to hydrolysis with weak alkali and establish the presence of this anhydride either by direct isolation or by reduction to diisobutyl piperazine. Such procedure would exclude the possibility of the secondary formation of this anhydride (99). Of course, it would be necessary to establish that enolic leucine anhydride is not attacked by alkali much more easily than the keto form investigated by Fischer and Schrauth. These considerations of the dependence of cleavage upon the constitution of the anhydride should be kept in mind when the problem is discussed, whether the acidity of the stomach or the alkalinity of the intestines are liable to effect the cleavage of dioxopiperazines.

Substitution of dioxopiperazines changes also their resistance to alkali. Thus O,O'-diacetyl-glycine- and -alanine anhydrides are not split by a boiling 5 N solution of alkali while the corresponding N,N'- compounds are completely hydrolyzed (100).

In reviewing the dioxopiperazine theory we may say that considerable experimental evidence has been adduced in its favor. On the other hand it was also shown how easily a secondary formation of dioxopiperazines takes place and how careful the investigator must be in the elimination of these possibilities. A modified dioxopiperazine theory based entirely on synthesis is suggested by Bergmann and his collaborators. This will be dealt with later in conjunction with the other theories which are built up on synthetic experiments.

We mentioned before that on the basis of researches into the polysaccharides it was assumed that associations and aggregations of elementary complexes are present in the proteins. However, this is to be regarded as a working hypothesis which cannot be generalized as yet although sporadic experiments are known in which simple compounds polymerize and the polymers show a behavior similar to that of proteins. The conception of proteins as aggregated complexes seemed to be so convincing that the existence of specifically acting disaggregating enzymes was first assumed by Oppenheimer (101). Pepsin was supposed to act in a disaggregating manner only by dissolving the subordinate valences which hold together the elementary complexes. It was regarded as a non-hydrolyzing ferment. This assumption could not be corroborated by experiment. It was also mentioned that a true valence linkage between anhydrides and amino acids was thought of as possible.

We here wish to discuss briefly the experiments which can be adduced in favor of the existence of associated compounds. It was shown first by Pfeiffer (102) that amino acids and polypeptides are capable of forming molecular compounds with neutral salts, e.g., $NaCl$, $CH_3 \cdot NH \cdot CH_2 \cdot COOH$, H_2O or $LiBr$, $NH_2 \cdot CH(CH_3)-COOH$, H_2O or $ZnCl_2$, $2NH_2 \cdot CH_2COOH$, $2H_2O$, etc. He also showed in collaboration with Angern (103) that not only proteins of high molecular weight but also amino acids are salted out easily, this behavior being independent from the solubility of the amino acid in water. Abderhalden and Sickel (104) observed the formation of mixed crystals of amino acids. Particularly important are the experiments by Pfeiffer and Angern

(105) which show that dioxopiperazines are capable of forming molecular combinations with amino acids, various salts and organic ompounds, e.g., glycine anhydride, $2LiCl$, $2.5H_2O$. Sarcosine anhydride gives molecular compounds with tryptophane scatole, anthranilic and p-aminobenzoic acids.

3. The experiments of Waldschmidt-Leitz

An extremely important criterion for the validity of an assumed structural formula is the enzyme test. It is well known that the enzymatic cleavage is very specific and depending upon fine configurational particulars. It indicates that a certain configuration but not that a particular compound occurs in nature (106). The attempts to split dioxopiperazines by enzymes were so far entirely unsuccessful (107). Levene and Meyer found that glycine anhydride is eliminated unchanged in the urine, in contrast to glycyl-glycine, the nitrogen of which is completely eliminated as urea. Abderhalden and Goto observed that a cleavage of 2,5-dioxopiperazines does not take place with pepsin. No enzymatic cleavage of 2,5-dioxopiperazines was observed by Waldschmidt-Leitz and Schäffner (108).

One may expect that the researches of Waldschmidt-Leitz and his collaborators on the separation of enzymes will contribute to the elucidation of the question of the protein structure.

It is well known that the foundation of Fischer's peptide theory, i.e., the conception that the CO·NH linkage is the characteristic linkage of proteins was strengthened particularly by the fact that synthetic peptides were split by proteolytic enzymes. The investigations of Willstätter and his collaborators on the purification and characterization of enzymes made a further progress in this direction possible.

In continuation of Willstätter's researches on the application of fractionated adsorption Waldschmidt-Leitz and Harteneck (109) succeeded in separating trypsin and erepsin the two proteolytic enzymes of the pancreas. Similarly a separation of trypsin and erepsin of the intestines could be effected (110). It could be proved that the two enzymes possess a highly specific action, since all simple dipeptides are split by erepsin of both the pan-

creas and the intestines, while proteins are not attacked by this enzyme. On the other hand trypsin attacks proteins. According to Waldschmidt-Leitz, the following specific enzyme groups must be distinguished (118): 1) erepsin, 2) trypsin not activated, 3) trypsin activated (trypsin + enterokinase), 4) pepsin.

The following table shows the differential behavior of erepsin and trypsin:

SUBSTITUTE	ENZYME			
	Erepsin from		Trypsin	
	Intestines	Pancreas	Activated	Not activated
Alanyl-tyrosine................................	+	+	—	—
Glycyl-tyrosine................................	+	+	—	—
Glycyl-glycine.................................	+	+	—	—
Leucyl-glycine.................................	+	+	—	—
Leucyl-alanine................................	+	+	—	—
Glycyl-alanine................................	+	+	—	—
Leucyl-glycyl-glycine..........................	+	+	—	—
Peptone (Merck)...............................	—	—	++	+
Clupein (herring)..............................	—	—	++	+
Thymus histone...............................	—	—	+	—
Casein...	—	—	+	—
Fibrin...	—	—	+	—
Gelatin..	—	—	+	—
Gliadin..	—	—	+	—
Zein...	—	—	+	—
Egg albumen...................................	—	—	+	—
Rhicinus globulin..............................	—	—	+	—

These results are in contrast with those of Fischer and Abderhalden (111) who distinguished between dipeptides which are hydrolyzed by enzymes and those which are not (112). According to Waldschmidt-Leitz, there is no such differentiation among the dipeptides, provided their configuration justifies the assumption that they occur in nature. Similarly all statements on the cleavage by erepsin of peptones, protamines, histones and casein must be revised. The complete hydrolysis of proteins is not feasible with a combination of two enzymes out of the four groups mentioned in contrast to older statements of Fischer and Abderhalden who assume that some enzyme groups may mutually

replace each other (113), or to the researches of Henriques and Gjaldbaek who find that an exhaustive treatment of casein with trypsin may replace the action of pepsin (114). Of course, the contrary results obtained by the older investigators are due to the impurity of enzymes used which represented mixtures with varying amounts of components. The action of enzymes is independent from the order of their application. E. Waldschmidt-Leitz found that the three enzyme groups, trypsin, trypsin + enterokinase, erepsin may be substituted for each other to a certain extent, although the enzymatic action of each is well defined. This was observed by following up the cleavage of casein effected by pepsin and erepsin-free trypsin. The progress of cleavage was estimated by the increase of the number of free carboxyl groups (115). The peptic action could be observed after the tryptic action had taken place. On the other hand when the hydrolysis is carried out with trypsin and erepsin combined, there is no substrate left which would make the peptic action visible. Under the action of pepsin and similarly under the action of non-activated trypsin, the protein molecule is split into large parts. This is indicated by a slight increase of the number of carboxyl groups. The activated trypsin on the other hand produces free amino acids from the protein. The cleavage of proteins by non-activated trypsin shows that trypsin does not occur in an inactive form as a zymogen which is activated by enterokinase. It is rather to be assumed that the activator enterokinase acts as an auxiliary substance for the cleavage of certain compounds (116). It was shown in a recent paper that the composition of the products of cleavage of clupein depends upon the sequence of the enzymes. But whatever enzyme is taken into consideration the cleavage is always one of the $CO \cdot NH$-linkage. In no case a disaggregation into elementary complexes was observed (117).

These conclusions are in a certain contrast with the results obtained by H. Steudel and collaborators (117a). If the action of a proteolytic enzyme always depends upon the dissolution of the regular $CO \cdot NH$-linkage, the amount of amino and carboxyl groups should always be the same, regardless of the method of

determination. However, when certain proteins (casein, serum globulin, serum albumin and gluten casein) were subjected to peptic cleavage the number of carboxyl groups formed was several times greater than that of the amino groups. It is impossible as yet to give a definite explanation of this phenomenon It might be caused by the cleavage of linkages different from the CO·NH-linkage, e.g., the ester linkage between a carboxyl group and the hydroxy group of a hydroxy amino acid. However, this would not account for the increase of amino nitrogen. Another explanation would be that groups are freed by peptic cleavage which are less basic than the α-amino group of the simple amino acids, since it is known that only a fraction of the nitrogen of some amino acids (e.g., tryptophane, arginine, lysine) is indicated by the method of Van Slyke.

Judging from the fact that the splitting action of individual enzymes stops at certain intermediate stages of the protein hydrolysis, one is justified in assuming that it will be possible by this method to realize the fractionated hydrolysis of proteins which will permit the isolation of biologically defined elementary complexes. Thus the investigations of Waldschmidt-Leitz corroborate the original idea of true valence linkages without leaving room for the conception of aggregates which are separated into smaller parts by enzymatic action. His objections to the dioxopiperazine theory are based on the same principles. He contends that altogether too much importance is attributed to the possibility of occurrence of dioxopiperazines (118). It is admitted that the anhydride structure may play a rôle in proteinoids like silk or certain skeleton constituents which are resistant to enzymes. But since no cleavage of 2,5-dioxopiperazines is observed under biological conditions (they also are very difficultly split at the pH of the gastro-intestinal tract) it is not to be assumed that they generally occur in proteins. Similar considerations hold for the other suggested cyclic structures, particularly those of Troensegaard and Bergmann.

It seems that this general statement goes too far, since it is based only on the action of enzymes on the simplest dioxopiperazines. As long as we do not have any information on the enzyme

action on the enolic tautomers of both dioxopiperazines and peptides, peptide-anhydride combinations and the other assumed ring structures, it is as yet unjustified to deny their occurrence in proteins. As the interest of the protein chemists is focussed on these researches, it is to be expected that a decision will be reached in the near future.

In this connection attention should be paid to the experiments of Levene and coworkers (119) on the action of enzymes on peptides and the circumstances influencing the formation of anhydrides.
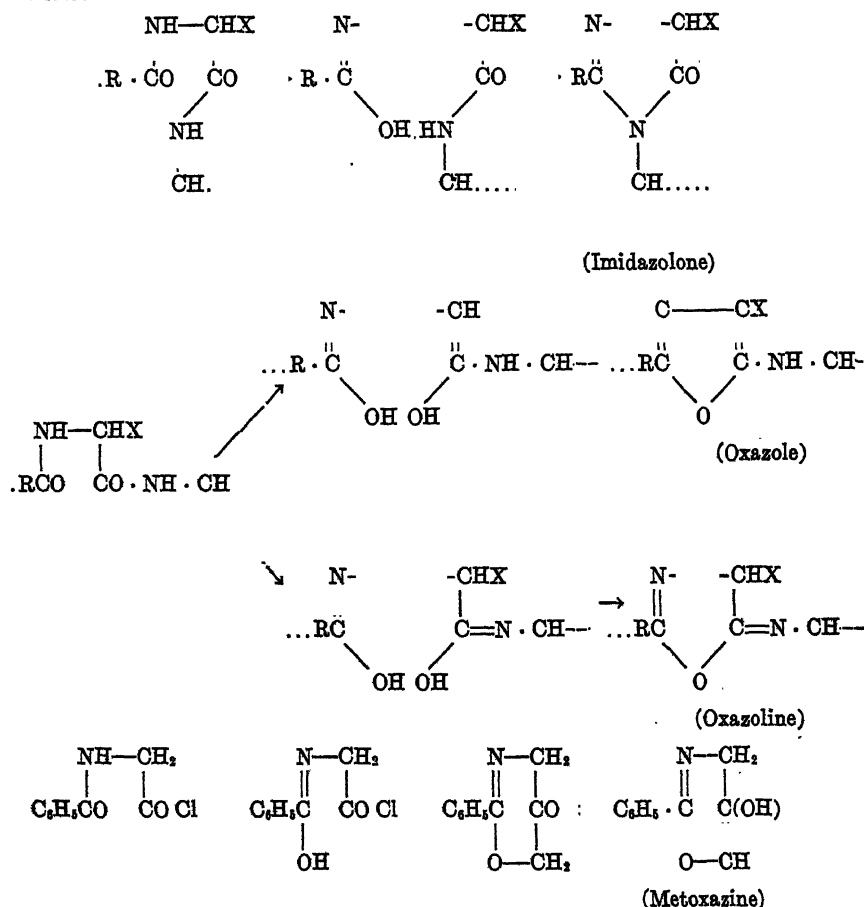
4. The synthetic heterocyclic compounds

We still have to deal with the interesting synthetic experiments of Karrer, Bergmann and their collaborators. These investigators attempt to approach the problem in a purely synthetic manner by preparing compounds of given structures which might possibly occur in proteins.

It was previously mentioned that Karrer first pointed to the possibility of existence of enolic forms of dioxopiperazines (70). The pertaining experiments were first carried out with dibenzyl dioxopiperazine derivatives, which were obtained by the action of benzyl chloride and ω-chloro-p-toluic acid ester respectively on the silver compound of glycine anhydride and have the formulas (120):
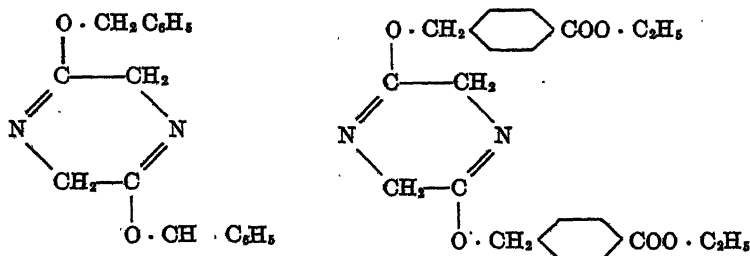

These compounds are particularly interesting with view to their extreme sensitiveness to dilute acids; dilute hydrochloric acid effects a cleavage with formation of glycine and benzyl chloride or the ω-chloro-p-toluic acid respectively. Karrer and coworkers contemplate also other possibilities of formation of anhydrides out of peptides, which are represented by the following equations.

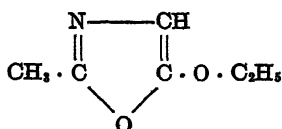They are characterized by the assumption of enolic rearrangement:

$$
\begin{array}{ccc}
\text{NH—CHX} & \text{N-\quad\quad-CHX} & \text{N-\quad\quad-CHX} \\
\underset{\text{NH}}{.\text{R}\cdot\overset{|}{\text{CO}}\quad\overset{|}{\text{CO}}} & \underset{\text{OH.HN}}{\cdot\text{R}\cdot\overset{\cdot\cdot}{\text{C}}\quad\overset{|}{\text{CO}}} & \underset{\text{N}}{\cdot\text{R}\overset{\cdot\cdot}{\text{C}}\quad\overset{|}{\text{CO}}} \\
\dot{\text{C}}\text{H.} & \underset{}{\text{CH}\ldots} & \text{CH}\ldots
\end{array}
$$

(Imidazolone)

$$
\begin{array}{cc}
\text{N-\quad\quad-CH} & \text{C———CX} \\
\underset{\text{OH OH}}{\ldots\text{R}\cdot\overset{\shortparallel}{\text{C}}\quad\overset{\shortparallel}{\text{C}}\cdot\text{NH}\cdot\text{CH--}} & \underset{\text{O}}{\ldots\text{R}\overset{\shortparallel}{\text{C}}\quad\overset{\shortparallel}{\text{C}}\cdot\text{NH}\cdot\text{CH-}}
\end{array}
$$

$$
\begin{array}{c}
\text{NH—CHX} \\
\underset{}{.\text{RCO}\quad\text{CO}\cdot\text{NH}\cdot\text{CH}}
\end{array}
$$

(Oxazole)

$$
\begin{array}{cc}
\text{N-\quad\quad-CHX} & \text{N-\quad\quad-CHX} \\
\underset{\text{OH OH}}{\ldots\text{R}\overset{\cdot\cdot}{\text{C}}\quad\overset{\|}{\text{C}}\text{=N}\cdot\text{CH--}} & \underset{\text{O}}{\ldots\text{R}\overset{\cdot\cdot}{\text{C}}\quad\overset{\|}{\text{C}}\text{=N}\cdot\text{CH--}}
\end{array}
$$

(Oxazoline)

$$
\begin{array}{cccc}
\text{NH—CH}_2 & \text{N—CH}_2 & \text{N—CH}_2 & \text{N—CH}_2 \\
C_6H_5CO\quad CO\,Cl & \underset{\text{OH}}{C_6H_5\overset{\|}{C}\quad CO\,Cl} & \underset{\text{O—CH}_2}{C_6H_5\overset{\|}{C}\quad\overset{|}{CO}} : & \underset{\text{O—CH}}{C_6H_5\cdot\overset{\|}{C}\quad C(OH)}
\end{array}
$$

(Metoxazine)

The formation of an imidazolone compound according to 1.) may be seen on the reaction of the hippuric acid amide with PCl$_5$. Here the 2-phenyl glyoxal-5-one is obtained:

$$
\begin{array}{c}
\text{N====CH}_2 \\
\underset{\text{NH}}{C_6H_5\cdot\overset{|}{C}\quad\overset{|}{CO}}
\end{array}
$$
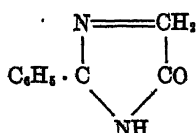
which is decomposed by acids with formation of hippuric acid.

The formation of oxazoles according to 2.) was observed in many cases. The formula of 2-methyl-5-ethoxyoxazole may serve as an example of the compounds of this group:

$$
\begin{array}{ccc}
\text{N} & \!\!\!\!\text{------} & \text{CH} \\
\| & & \| \\
\text{CH}_3 \cdot \text{C} & & \text{C} \cdot \text{O} \cdot \text{C}_2\text{H}_5 \\
& \diagdown \diagup & \\
& \text{O} &
\end{array}
$$

Phenylhydroxydihydrometoxazine (3) is obtained by the action of diazomethane on hippuryl chloride.

The imidazolones, oxazoles and metoxazines are likewise extremely sensitive to acids. In some cases an acid concentration which is necessary for the activation of pepsin is sufficient to effect the cleavage. If these compounds were built entirely of amino acids, this treatment would lead to peptides. The formation of the ring takes place the more easily the lower the fatty acid, which is combined with the amino acid (121).

The results obtained by Karrer are no doubt extremely interesting, particularly the observation that compounds related to peptides may be obtained by the cleavage of certain ring structures at an extremely slight acid reaction, such as is found in the stomach. Since these researches are not yet directly applicable to the conditions prevailing in natural proteins, it is very desirable that attempts be made to synthetize these ring structures entirely out of amino acids and to study these under conditions which prevail in the gastro-intestinal tract.
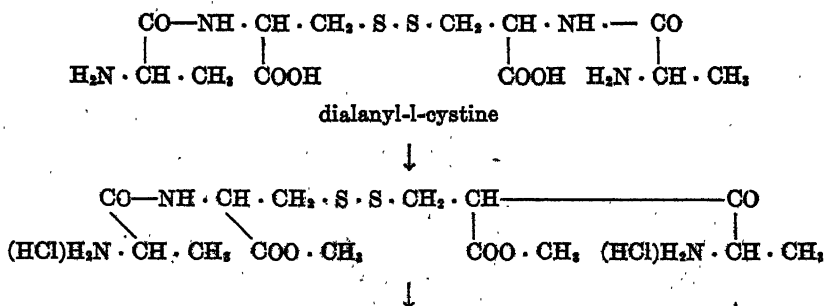
5. The iso- and allodioxopiperazines

A series of important synthetic investigations was carried out by Bergmann and his collaborators. In continuation of the studies on the rearrangement of compounds which show a behavior similar to peptides they succeeded in isolating compounds which were regarded first as oxazoline derivatives. Later they were recognized as dioxopiperazine derivatives. In contrast to the other dioxopiperazines these compounds show a tendency to polymerization (122). Bergmann showed in collaboration with
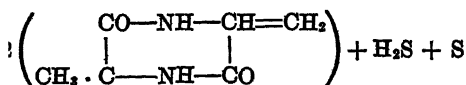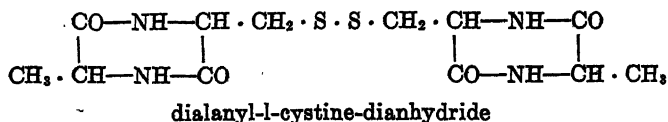
Miekeley and Kann that these compounds are formed from ester chlorides of the dipeptides glycyl-and alanyl-serine on treatment with thionyl chloride. They show a peculiar behavior on treatment with acids and alkalis. With acid one-half of their nitrogen content is split off as ammonia with formation of pyruvic acid. On treatment with alkali and subsequent neutralization polymerization takes place with formation of compounds which are insoluble in water and other solvents. The series of reactions which are involved in these changes may be illustrated by the following equations showing the changes to which alanyl serine may be subjected:

6-methyl-2, 3, 5-trioxopiperazine

pyruvic acid

Disodium compound

$(C_6H_8O_2N_2)x$

Tetrapeptide
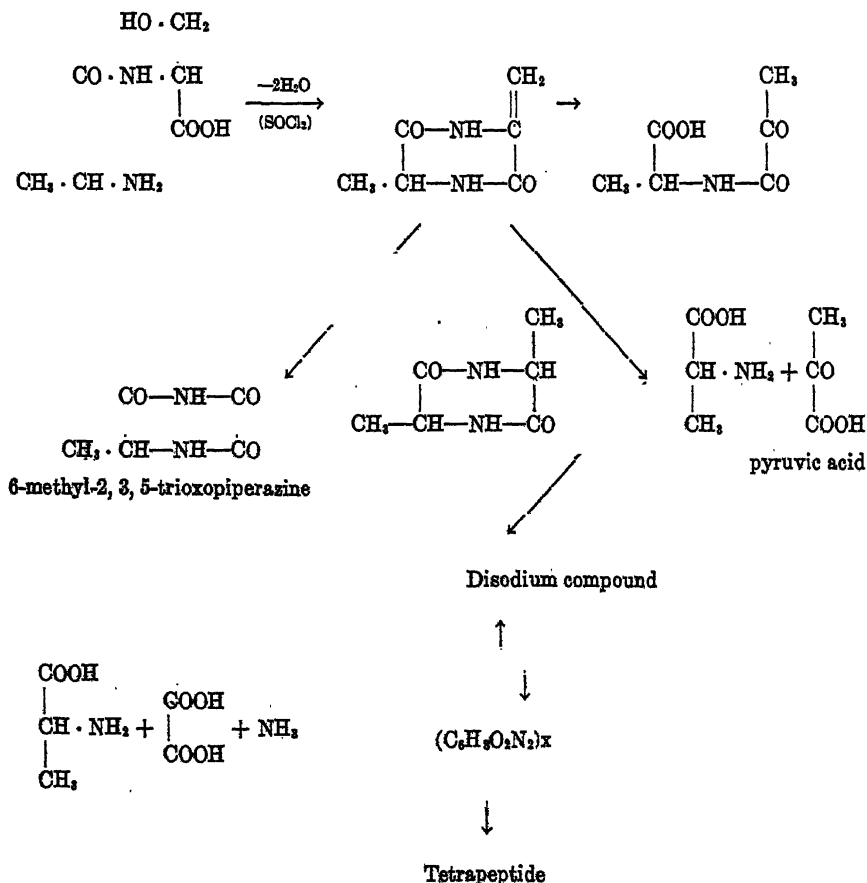
· This scheme of reactions shows the following: The catalytic hydrogenation which leads to alanine anhydride proves that the assumed formula of a methylene dioxopiperazine is correct. It also shows that there is a clear relation to pyruvic acid to which an important biological rôle is attributed. Solution in alkali leads to a yellow compound from which a polymer compound is precipitated with acid. This reaction is reversible. The polymer compound yields on acid hydrolysis a tetrapeptide. This would suggest the possibility of four amino acids constituting this compound without any piperazine nuclei. However, this assumption cannot be made since this same polymer gives on reduction alanine anhydride, i.e., a compound with two amino acids only. The determination of molecular weight in phenol corresponds to the formula $(C_6H_8O_2N_2)_2$ (mol. weight 280). However, this is not necessarily the true molecular weight, but possibly that of elementary complexes disrupted by the action of phenol.

Bergmann and Stather (123) obtained the same compound also in a different way. They started with cystine which according to known methods was transformed into dialanyl-l-cystine and the corresponding diester. From this the new dialanyl-l-cystine dianhydride was obtained. This compound is insoluble in water, but dissolves quickly on addition of alkali. When the solution is neutralized with acid a crystalline product is obtained and simultaneously sulfur and hydrogen sulfide form. The crystalline compound has the composition of a polymer 3-methylene-6-methyl-2,5-dioxopiperazine which is obtained from alanyl serine as described before. The reactions which lead to the formation of this compound are:

$$CO-NH \cdot CH \cdot CH_2 \cdot S \cdot S \cdot CH_2 \cdot CH \cdot NH \cdot - CO$$
$$| \quad \quad | \quad \quad \quad \quad \quad | \quad \quad |$$
$$H_2N \cdot CH \cdot CH_3 \quad COOH \quad \quad COOH \quad H_2N \cdot CH \cdot CH_3$$

dialanyl-l-cystine

↓

$$CO-NH \cdot CH \cdot CH_2 \cdot S \cdot S \cdot CH_2 \cdot CH \text{———————} CO$$
$$\diagdown \quad \quad \diagdown \quad \quad \quad \quad \quad | \quad \quad |$$
$$(HCl)H_2N \cdot CH \cdot CH_3 \quad COO \cdot CH_3 \quad \quad COO \cdot CH_3 \quad (HCl)H_2N \cdot CH \cdot CH_3$$

↓

$$\begin{array}{l} \cdot CO-NH-CH \cdot CH_2 \cdot S \cdot S \cdot CH_2 \cdot CH-NH-CO \\ \quad\quad\quad | \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad | \\ CH_3 \cdot CH-NH-CO \quad\quad\quad\quad\quad CO-NH-CH \cdot CH_3 \end{array}$$

dialanyl-l-cystine-dianhydride

↓

$$\left(\begin{array}{l} \quad\quad CO-NH-CH{=}CH_2 \\ \quad\quad\;\; | \quad\quad\quad\quad\quad | \\ CH_3 \cdot C----NH-CO \end{array} \right) + H_2S + S$$

↓

$(C_6H_8O_2N_2)x$

It is remarkable that the formation of the two anhydrides in cystine so increases the lability of the sulfur atoms that 71 per cent of the sulfur of that compound is split off, while the $-C-S-$ linkage in dialanyl cystine remains entirely intact under the same conditions. It seems that this reaction is not limited to serine and cystine derivatives but is shown as well by other amino acids containing no hydroxyl or sulfur groups.

In addition to the polymer isodioxopiperazines, another polymer modification has been prepared recently by Bergmann, Miekeley and Kann (123a). They found that it is sufficient to heat methylenemethyldioxopiperazine with a dilute aqueous solution of arginine, to obtain the "allo-form" which is different from the "iso-form" previously obtained. Other weak bases (ammonia, guanidine, etc.) act similarly. Treating with concentrated hydrochloric acid leads to a tetrapeptide which is identical with that obtained from the "iso-form."

$$\begin{array}{l} \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\nearrow \quad [C_6H_8O_2N_2]x \\ \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\;\; \text{Isomethylenemethyldioxopiperazine} \\ CO-NH-C{=}CH_2 \quad\quad / \\ \\ CH_3 \cdot CH-NH-CO \quad\quad \searrow \\ \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\searrow \quad [C_6H_8O_2N_2]y \\ \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\;\; \text{Allomethylenemethyldioxopiperazine} \end{array}$$
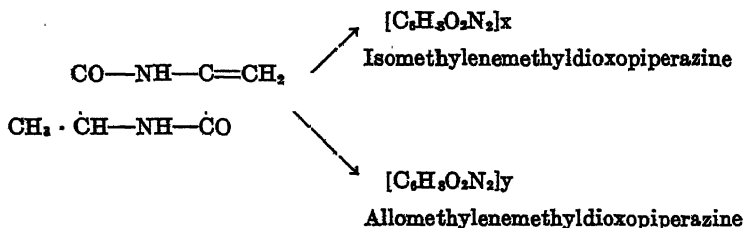
Particular importance is attributed to this finding, since it is shown for the first time that the anhydride of two amino acids

which frequently occur in native proteins may be transformed into an isomer form in aqueous solution by an amino acid which likewise occurs in natural proteins. This in turn permits the conclusion that the allodioxopiperazines themselves occur in natural proteins.

The allo-form is chemically very interesting. Upon treatment with formaldehyde a compound results to which the designation monoformal-methylenemethyldioxopiperazine is given. This compound swells in the presence of water, forms a jelly, causes no noticeable depression of the freezing temperature and gives on drying in thin layers films which may be made sensitive to light on treatment with chromate just like gelatin films. It is concluded that compounds which show a colloidal behavior similar to that of gelatin, do not necessarily have to consist of a great number of different amino acids.

In contrast to the ordinary dioxopiperazines (and dibenzalpiperazine) which bind only a small quantity of tannin, the iso- and allodioxopiperazines possess an outspoken capacity to absorb both tannin and dyestuffs (malachite green, acid fuchsin) similar to some natural proteins. The following table illustrates this behavior. The results were obtained by the colorimetric comparison of solutions originally containing a known quantity of malachite green which were shaken with a given quantity of the compound.

Compound	Adsorption per cent
Glycine anhydride	0.9
Methylenedioxopiperazine	2.2
Isomethylenedioxopiperazine	35.6
Allomethylenedioxopiperazine	16.7
Methylenemethyldioxopiperazine	6.3
Isomethylenemethyldioxopiperazine	20
Allomethylenemethyldioxopiperazine	60.3
Silk fibroin	100
Zephir wool	97.6
Benzaldehyde compound of methylenedioxopiperazine	98.3

This "super molecular" state of certain dioxopiperazines seems to be related to the state of some natural proteins (gelatin, silk fibroin gliadin). Also the results of the determination of molec-

ular weight of these proteins (200–400) would fit into this conception. Like proteins the iso- and allodioxopiperazines bind dyestuffs and tannins. It is emphasized that the stabiliy of these substances depends upon the presence of the methylene group, while the stability of dioxopiperazines of the proteins in case these are actually present, is probably brought about in a different manner.
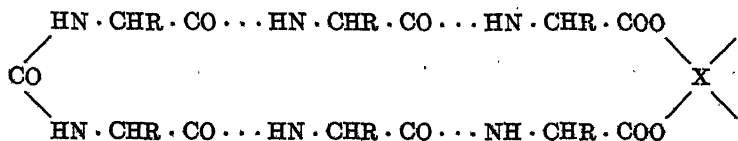
Bergmann's investigations lead to a new conception of the structure of the compounds of high molecular weight (124). The following comparison is used to illustrate these conditions. When a compound enters into the crystallized state, the individual molecules cease to exist since they are now connected with each other by lattice forces. Thus a crystal is to be regarded as a kind of supermolecular structure, it being possible to restore the molecular subdivision through destruction of the crystallized state by vaporization, solution or liquefaction. This supermolecular state is not a structural constant but a form of a state which is determined by the surrounding physical and chemical conditions. Similarly the conclusion that a protein appears to be of a high molecular weight is drawn as a rule from its behavior toward water, although it may be "molecularly disperse" in phenol. Thus the forces which bring about this phenomenon are comparable to the lattice forces, the difference being only a quantitative one. On the other hand, the hydrogenation of the polymer isodioxopiperazines leads to compounds with six carbon atoms while acid hydrolysis and the determinations of molecular weight of the same substance in phenol point to the existence of compounds with twelve carbon atoms. This fact would suggest that there is no sharply defined elementary complex which would first come into appearance in all degradations. This seems to be characteristic of all proteins. There is another resemblance between these polymer isodioxopiperazines and the proteins. The fact that polypeptides result on hydrolysis of proteins makes the assumption possible that the degradation does not lead over the dioxopiperazines first. It is thinkable that the disaggregation by ferments leads directly to polypeptides which are further

split by another enzyme group, just as the acid hydrolysis of the isodioxopiperazines leads to tetrapeptides.
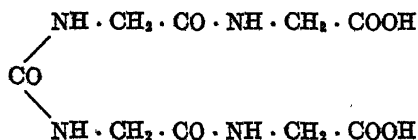
This interesting theory of Bergmann also requires biological corroboration. In its present formulation it would make appear the protein problem clear and reduced to comparatively simple basic ideas.

6. The ureide theory

Finally the ureide linkage is to be mentioned, the occurrence of which is suggested by Brigl and Held (53). These authors base their assumption on the fact that neither the pure peptide nor the dioxopiperazine theories account for the relation of oxygen to nitrogen being in most proteins considerably higher than 1, while according to either theory it should be approximately 1. Therefore, the hypothesis is advanced that proteins contain polypeptide chains of varying lengths which are connected with each other by means of groups containing oxygen. It is possible that they possess ureide structure according to the schematic formula:

$$
\begin{array}{c}
\text{HN} \cdot \text{CHR} \cdot \text{CO} \cdots \text{HN} \cdot \text{CHR} \cdot \text{CO} \cdots \text{HN} \cdot \text{CHR} \cdot \text{COO} \\
/ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \backslash \ / \\
\text{CO} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{X} \\
\backslash \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad / \ \backslash \\
\text{HN} \cdot \text{CHR} \cdot \text{CO} \cdots \text{HN} \cdot \text{CHR} \cdot \text{CO} \cdots \text{NH} \cdot \text{CHR} \cdot \text{COO}
\end{array}
$$

The possibility of an ureide structure was discussed repeatedly in the chemical literature (125). Brigl and Held used as a model for their investigation the diureide of the dipeptide glycyl glycine:

$$
\begin{array}{c}
\text{NH} \cdot \text{CH}_2 \cdot \text{CO} \cdot \text{NH} \cdot \text{CH}_2 \cdot \text{COOH} \\
/ \\
\text{CO} \\
\backslash \\
\text{NH} \cdot \text{CH}_2 \cdot \text{CO} \cdot \text{NH} \cdot \text{CH}_2 \cdot \text{COOH}
\end{array}
$$

The results obtained by fusion with phthalic anhydride, which was suggested by Brigl and Klenk (126) as a possible reagent for the isolation of definite compounds out of proteins and by the application of the previously mentioned hypobromite method of

Goldschmidt and Steigerwald do not disagree with the assumption of the ureide structure of proteins. The experiments with enzymes (pepsin and trypsin), however, gave entirely negative results. It is possible that ureides of other peptides will behave differently.

It is not desired to carry further our presentation of the researches into the structure of proteins. Although it was intended to illustrate the different phases and branches of development as completely as possible this paper does not by any means contain a complete review of all the work that has been done in this realm.

It is obvious that the research into the structure of the proteins is in a constant flux at the present time. There are radical theories which regard even amino acids as secondary products and there are others which view amino acids and polypeptides as elementary compounds related directly to the various cyclic structures. It is impossible at the present time to unconditionally accept any one of the theories described. They all have their weak and strong points. It is possible that the near future will bring fundamental developments in this realm. New facts may become known to which one or another conception may fall prey. Nevertheless a compilation of pertaining facts and ideas at this, as it seems, critical point, may involve psychological and educational interest, since it allows an easier contraposition of the state of scientific progress in given periods.

In considering the possibilities of future development, we believe that the investigator who tries to elucidate the riddles of protein structure is entitled to the same optimism which guided Emil Fischer in his work, as pronounced in a statement, part of which may be given in free translation:

Nature accomplished her highest achievement in building up proteins and their various derivatives. The belief that she confined herself to only a few types would disagree with all experiences of chemistry and biology. . . . .

If by a lucky accident it should become possible today to prepare a

genuine protein by some brutal reaction, e.g., by fusing amino acids in the presence of a dehydrating agent, and if it were possible to identify this artificial product with a natural one, only little would have been achieved for the chemistry and practically nothing for the biology of proteins. . . . . I should feel inclined to regard it as good luck, that chemistry is forced to create numerous new methods of synthesis, idenfication and isolation and to closely study hundreds of intermediate products, before it solves the protein problem. For ultimately these methods will not only serve in the preparation of all proteins of nature and of more than nature created. They will probably suffice to elucidate the numerous and remarkable products of transformation of proteins, which play such an important rôle as enzymes, toxins, and others.

REFERENCES

(1) E. FISCHER, Z. angew. Chem. **20**, 913 (1907); Z. physiol. Chem. **99**, 54 (1907).

(2) E. ABDERHALDEN, Ber. **49**, 561, (1916).

(3) F. BOPP, Ann. **69**, 28 (1849).

(4) E. ERLENMEYER, Ann. **119**, 17 (1861).

(5) H. HLASIVETZ AND J. HABERMANN, Ann. **159**, 328 (1871).

(6) R. COHN, Ber. **29**, 1786 (1896).

(7) E. ABDERHALDEN AND E. FISCHER, Ber. **39**, 753, 2315 (1906); **40**, 3555 (1907).
E. ABDERHALDEN AND SUWA, Z. physiol. Chem. **66**, 13 (1910).

(8) H. D. DAKIN, Biochem. J. **12**, 290 (1918).

(9) H. D. DAKIN, J. Biol. Chem. **44**, 499 (1920).

(10) E. FISCHER, Ber. **38**, 607 (1905); **39**, 607 (1906).
E. ABDERHALDEN, Lehrb. d. physiol. Chem. 3rd edition, 885 (1915).

(11) P. A. LEVENE AND F. J. BIRCHARD, J. Biol. Chem. **13**, 277 (1912).
P. A. LEVENE AND J. V. D. SCHEER, J. Biol. Chem., **22**, 425 (1915).

(12) A. KOSSEL, Ber. **34**, 3214 (1901).

(13) O. COHNHEIM, Chemie der Eiweisskörper, 3rd edition **77**, (1911).
R. O. HERZOG, Eiweisstoffe in Ullmann's Enzyclopädie **4**, 494 (1916).
E. STIASNY, Collegium (1920) 256. Science **54**, 483 (1923).
K. HESS, Z. Elektrochem. **26**, 232 (1920).

(14) W. S. SSADIKOV AND N. D. ZELINSKY, Biochem. Z. **136**, 241, (1923); **147**, 30 (1924).

(15) N. TROENSEGAARD, Z. physiol. Chem. **122**, 87 (1920); **127**, 137 (1923).
N. TROENSEGAARD AND EUG. FISCHER, Z. physiol. Chem. **142**, 35 (1925); **143**, 304 (1925).
N. TROENSEGAARD, Z. angew. Chem. **38**, 623 (1925).
N. TROENSEGAARD AND B. KOUDAHL, J. physiol. Chem. **153**, 93 (1926).

(16) H. D. DAKIN, J. Biol. Chem. **44**, 499, (1920).

(17) E. KOMM, Z. physiol. Chem. **156**, 161 (1926).

(18) N. TROENSEGAARD AND EUG. FISCHER, Z. physiol. Chem. **143**, 304 (1925).

(19) N. TROENSEGAARD, Z. physiol. Chem. **134**, 100 (1924).

(20) TH. CURTIUS, Ber. **16**, 753 (1883); **17**, 953 (1884).

(21) F. W. ABENIUS AND O. WIDMANN, Ber. 21, 1663 (1883); J. prakt. Chem. 38, 299, 305 (1888).
C. A. BISCHOFF, Ber. 21, 1257 (1888); 25, 2951 (1892).
F. W. ABENIUS, J. prakt. Chem. [2] 40, 425 (1889).

(22) E. ABDERHALDEN AND H. QUAST, Z. physiol. Chem. 151, 146 (1926).

(23) F. W. ABENIUS. J. prakt. Chem. [2] 41, 80 (1890).
C. A. BISCHOFF, Ber. 25, 2954 (1892).

(24) P. GLAGOLEV, Biochem. Z. 70, 119 (1914).

(25) F. MYLIUS, Ber. 17, 292 (1884).

(26) E. HOYER, Z. physiol. Chem. 34, 347 (1902).

(27) R. COHN, Z. physiol. Chem. 22, 170, (1897); 29, 283 (1900).

(28) E. ABDERHALDEN AND K. FUNK, Z. physiol. Chem. 53, 19 (1907).

(29) R. BRILL, Ann. 434, 204 (1923).

(30) R. O. HERZOG AND M. KOBEL, Z. physiol. Chem. 134, 299, (1923).

(31) E. ABDERHALDEN AND W. STIX, Z. physiol. Chem. 129, 143 (1923); 132, 238 (1923).

(32) S. P. L. SØRENSON, J. Chem. Soc. 2995 (1926).

(33) S. S. GRAVE, T. W. MARSHALL AND H. W. ECKWEILLER, J. Am. Chem. Soc. 39, 112 (1917).

(34) P. BRIGL, Ber. 56, 1887 (1923).

(35) E. ABDERHALDEN AND E. KOMM, Z. physiol. Chem. 134, 121 (1924); 139, 121 (1924).

(36) A. KOHLER, Ann. 134, 369 (1865).

(37) S. SALASKIN, Z. physiol. Chem. 32, 592 (1901).
S. SALASKIN AND K. KOWALEWSKY, Z. physiol. Chem. 38, 573 (1903).

(38) E. ABDERHALDEN AND E. KOMM, Z. physiol. Chem. 145, 308, (1925).

(39) P. A. LEVENE AND G. B. WALLACE, Z. physiol. Chem. 47, 143 (1906).
P. A. LEVENE AND W. A. BEATTY, Ber. 39, 2060 (1906).

(40) E. ABDERHALDEN, Z. physiol. Chem. 154, 18 (1926).

(41) E. ABDERHALDEN AND E. ROSSNER, Z. physiol. Chem. 152, 271 (1926).

(41a) S. FRÄNKEL, Biochem. Z. 120, 218; 130, 592; 134, 308; 145, 225 (1921-4).

(42) E. ABDERHALDDN, Z. physiol. Chem. 131, 284 (1923).

(43) E. FISCHER, Ber. 35, 1164 (1908).

(44) E. ABDERHALDEN, Z. physiol. Chem. 131, 281, 284 (1923).
E. ABDERHALDEN AND E. KOMM, Z. physiol. Chem. 132, 1 (1924); 134, 113, (1924); 136, 134 (1924).

(45) E. ABDERHALDEN AND W. STIX, Z. physiol. Chem. 129, 143, (1923).

(46) E. ABDERHALDEN AND K. FUNK, Z. physiol. Chem. 64, 436, (1910).

(47) M. LÜDTKE, Z. physiol. Chem. 151, 215 (1925).

(48) C. PAAL, Ber. 27, 974 (1894).
B. KÜHN, Ber. 17, 2880 (1884).

(49) M. BERGMANN AND L. ZERVAS, Z. physiol. Chem. 152, 282 (1916).

(50) E. ABDERHALDEN AND W. STIX, Z. physiol. Chem. 132, 238 (1925).

(51) E. ABDERHALDEN, E. KLARMANN AND E. SCHWAB, Z. physiol. Chem. 134, 180 (1924).

(52) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. 139, 169 (1924).

(53) P. BRIGL AND H. HELD, Z. physiol. Chem. 152, 230 (1926).

(54) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **139**, 68 (1924), **143**, 290 (1925).
(55) O. LOEW, J. prakt. Chem. [2] **31**, 129 (1885).
(56) FR. KUTSCHER AND M. SCHENK, Ber. **38**, 455 (1905).
(57) E. ABDERHALDEN, E. KLARMANN AND E. KOMM, Z. physiol. Chem. **140**, 92 (1924).
(58) E. ABDERHALDEN AND E. KOMM, Z. physiol. Chem. **143**, 128 (1925).
(59) S. GODSCHMIDT AND CH. STEIGERWALD, Ber. **58**, 1346 (1925).
(60) Z. STARY, Z. physiol. Chem. **144**, 147 (1925).
(61) M. JAFFE, Z. Physiol. Chem. **10**, 399 (1886).
(62) T. SASAKI, Biochem. Z. **114**, 63 (1921).
(63) H. A. DOX, Z. physiol. Chem. **150**, 118 (1926).
(64) E. ADERHALDEN AND E. KOMM, Z. physiol. Chem. **139**, 181 (1924); **140**, 99 (1924).
    E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **148**, 20 (1925).
(65) E. BRAND AND M. SANDBERG, J. Biol. Chem. **70**, 381 (1926).
(66) E. ABDERHALDEN AND H. SUSUKI, Z. physiol. Chem. **127**, 281 (1923); **129**, 106 (1923).
(67) E. ABDERHALDEN, Z. physiol. Chem. **131**, 284 (1923).
(68) E. ABDERHALDEN AND E. KOMM, Z. physiol. Chem. **134**, 113, (1924); **136**, 134 (1924).
(69) E. ABDERHALDEN AND KLARMANN, Z. physiol. Chem. **135**, 199 (1924).
(70) P. KARRER, CH. GRÄNACHER AND A. SCHLOSSER, Helv. Chim. Acta **6**, 1108 (1923).
(71) E. ABDERHALDEN AND E. KLARMANN, Z. physiol. Chem. **139**, 64 (1924).
    E. ABDERHALDEN, Naturwissenschaften **13**, 99 (1925).
(72) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **149**, 100 (1925).
(73) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **149**, 298 (1925).
(74) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **152**, 90, (1926).
(75) E. ABDERHALDEN AND F. GEBELEIN, Z. physiol. Chem. **152**, 125 (1926).
(76) E. ABDERHALDEN AND R. HAAS, Z. physiol. Chem. **155**, 195, (1926).
(77) E. FISCHER, Ber. **38**, 2914 (1905).
(78) E. ABDERHALDEN AND R. HAAS, Z. physiol. Chem. **155**, 202 (1926).
(79) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **153**, 83 (1926).
(80) W. STENSTRÖM AND M. REINHARD, J. Biol. Chem. **66**, 819, (1925).
(81) V. SHIBATA AND T. ASAHINA, Bull. Chem. Soc. Japan **1**, 71 (1926).
(82) J. N. HARTLEY, J. Chem. Soc. **39**, 153 (1881); **41**, 481 (1882); Proc. Chem. Soc. (1889), 46.
(83) L. BALBIANO AND D. TRASCIATI, Ber. **33**, 2323 (1900), **34**, 1501 (1901).
(84) L. C. MAILLARD, C. r. **153**, 1078 (1911); Ann. chim. phys. [9], **1**, 519, 2210 (1914).
(85) E. ABDERHALDEN AND R. HAAS, Z. physiol. Chem. **155**, 197 (1926).
(86) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **154**, 99 (1926).
(87) E. ABDERHALDEN AND E. KLARMANN, Z. physiol. Chem. **139**, 64 (1924).
(88) E. ABDERHALDEN AND R. HAAS, Z. physiol. Chem. **153**, 147, (1926).
(89) KAITO SHIBATA, Acta Phytochem. **2**, 39 (1925).
(90) E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **149**, 299 (1925).
    E. ABDERHALDEN AND E. SCHWAB, Z. physiol. Chem. **148**, 254 (1925).

(91) E. ABDERHALDEN AND E. SCHWAB, J. physiol. Chem. **158**, 66 (1926).
(92) E. ABDERHALDEN AND E. ROSSNER, Z. physiol. Chem. **144**, 219 (1925).
(93) E. ABDERHALDEN AND KOHL-EGGER, Z. physiol. Chem. **156**, 128 (1926).
(94) E. ABDERHALDEN AND R. HAAS, Z. physiol. Chem. **151**, 114 (1926).
(95) M. LÜDTKE, Z. physiol. Chem. **141**, 100 (1924).
(96) E. ABDERHALDEN AND R. HAAS, Z. physiol. Chem. **155**, 202, (1926).
(97) E. ABDERHALDEN AND H. QUAST, Z. physiol. Chem. **151**, 145 (1926)..
(98) E. FISCHER AND W. SCHRAUTH, Ann. **354**, 21 (1907).
(99) P. BRIGL AND R. HELD, J. physiol. Chem. **152**, 233 (1926).
(100) E. ABDERHALDEN AND E. KOMM, Z. physiol. Chem. **140**, 108 (1924).
(101) K. OPPENHEIMER, Die Fermente and ihre Wirkungen, vol. 2, 811, 1208 (1926).
(102) P. PFEIFFER, Z. physiol. Chem. **133**, 22 (1924).
(103) P. PFEIFFER AND O. ANGERN, Z. physiol. Chem. **134**, 180, (1924).
(104) E. ABDERHALDEN AND H. SICKEL, Z. physiol. Chem. **134**, 29 (1924).
(105) P. PFEIFFER AND O. ANGERN, Z. physiol. Chem. **143**, 265, (1925).
(106) E. ABDERHALDEN, Z. physiol. Chem. **130**, 205 (1925).
(107) P. A. LEVENE AND G. M. MEYER, Am. J. Physiol. **25**, 214 (1909/10).
     E. ABDERHALDEN AND K. GOTO, Fermentforschung 7, 169, (1923).
(108) E. WALDSCHMIDT-LEITZ AND SCHÄFFNER, Ber. **58**, 1356 (1925).
(109) E. WALDSCHMIDT-LEITZ AND A. HARTENECK, Z. physiol. Chem. **147**, 286 (1925); **149**, 203 (1925).
(110) E. WALDSCHMIDT-LEITZ AND A. SCHÄFFNER, Z. physiol. Chem. **151**, 34 (1926).
(111) E. FISCHER AND E. ABDERHALDEN, J. physiol. Chem. **46**, 52 (1905).
(112) E. ABDERHALDEN, Z. physiol. Chem. **151**, 151 (1926).
(113) E. FISCHER AND E. ABDERHALDEN, Z. physiol. Chem. **39**, 81; **40**, 215 (1903).
(114) V. HENRIQUES AND J. K. GJALDBAEK, Z. physiol. Chem. **75**, 363 (1911); **83**, 83 (1912).
(115) R. WILLSTÄTTER AND E. WALDSCHMIDT-LEITZ, Ber. **54**, 2988 (1921); **55**, 3601 (1922).
(116) E. WALDSCHMIDT-LEITZ, Z. physiol. Chem. **132**, 181 (1923).
(117) E. WALDSCHMIDT-LEITZ AND E. SIMMONS, Z. physiol. Chem. **156**, 117 (1926).
(118) E. WALDSCHMIDT-LEITZ, Naturwissenschaften **14**, 129, (1926).
     E. WALDSCHMIDT-LEITZ, A. SCHÄFFNER AND W. GRASSMANN, Z. physiol. Chem. **156**, 74 (1926).
(119) P. A. LEVENE, H. S. SIMMS AND M. H. PFALTZ, J. Biol. Chem. **61**, 445 (1923); **62**, 711 (1924); **70**, 253 (1926).
(120) P. KARRER AND CH. GRÄNACHER, Helv. Chim. Acta 7, 763 (1924).
     P. KARRER AND R. WIDMER, Helv. Chim. Acta 8, 203 (1925).
(121) P. KARRER, E. MIYIAMICHI, H. C. STORM, AND R. WIDMER, Helv. Chim. Acta 8, 205 (1925).
(122) M. BERGMANN, A. MIEKELEY AND E. KANN, Z. physiol. Chem. **140**, 128 (1924).
     M. BERGMANN, A. MIEKELEY, FR. WEINMANN, AND E. KANN, Z. physiol. Chem. **143**, 108 (1925).
     M. BERGMANN, Ann. **445**, 17 (1925).
     M. BERGMANN, A. MIEKELEY, AND E. KANN, Z. physiol. Chem. **146**, 247 (1925).

(123) M. BERGMANN AND F. STATHER, Z. physiol. Chem. **152**, 189 (1926).
(124) M. BERGMANN, Naturwissenschaften **13**, 1045 (1925). Koll. Z. **40**, 289, (1926).
(125) SCHÜTZENBERGER, Bull soc. chim. **23**, 193 (1875).
       T. HABERMANN AND R. EHRENFELD, Z. physiol. Chem. **30**, 453 (1900)
       F. LIPPICH, Z. physiol. Chem. **90**, 441 (1914).
(126) P. BRIGL AND E. KLENK, Z. physiol. Chem. **131**, 83 (1923).

PHOSGENE

G. MALCOLM DYSON

Laboratory of Applied Pathology and Preventive Medicine, Manchester, England

The aim of this monograph is to provide a summary of the work done with phosgene. The literature of any substance whose chief use is in the capacity of a reagent, is hard to search, since the name of the reagent does not appear in the title of the paper, nor, as a rule in the index of the journal in which the paper is published; and bearing this in mind it is not claimed that the present work embodies all the references to phosgene which have been made in the literature. However, an attempt has been made to include all the important work done on the subject up to the end of 1925.

HISTORY AND PREPARATION OF PHOSGENE

During the years 1800–1812 several investigators, including Gay-Lussac and Thenard,[1] and Murphy,[2] published statements to the effect that carbon monoxide and chlorine were mutually inactive, but in 1812 John Davy[3] (brother of Humphrey) more carefully investigated their interaction. He mixed equal volumes of the two gases, which had been dried over calcium chloride, and exposed the mixture, contained in a vessel over mercury, to sunlight. In fifteen minutes the colour of the chlorine had entirely disappeared and a new gas which was possessed of lachrymatory properties remained. This Davy christened "phosgene" ($\varphi\tilde{\omega}s$ "light," and $\gamma\epsilon\nu\nu\acute{a}\omega$ "I give rise to") while he noted that the new gas did not fume in air, but reddened blue litmus and reacted with ammonia to give ammonium chloride. Further, he noticed that it was chemically very reactive and was decomposed by heating

[1] Gay Lussac and Thenard. Recherches Physico-chimique, Tom. II, 150.
[2] Nicholson's Journal, XXX, 227.
[3] Phil. Trans. (1812), p. 144.

with zinc, antimony or arsenic, and that it combined readily on warming with the oxides of certain elements to give the metal chloride and carbon dioxide. He showed, also, that no combination of carbon monoxide and chlorine took place in the absence of sunlight, even on passing through an earthenware tube heated to redness.

Several investigators carried on the work with phosgene which they obtained by various modifications of Davy's original method. Thus, for instance, Wilm and Wischin[4] (1868) obtained their phosgene by passing carbon monoxide, which had been purified by passage through caustic soda solution, and milk of lime, mixed with a slight excess of chlorine, which had been dried over strong sulfuric acid, into ten litre glass globes, exposed to sunlight. The phosgene was then absorbed in alcohol to give chlorocarbonic ester, of which these investigators state that they were able to prepare about 2.3 pounds a day.

Prior to 1860 only three methods (besides that of Davy), and these of little importance, had been recorded for the preparation of phosgene. They included the action of a mixture of concentrated nitric and hydrochloric acids on carbon bisulfide, discovered by Berzelius and Marcet,[5] and the decomposition of hexachloromethyl oxalate by heat (Cahours)[6] which takes place thus:

$$\begin{array}{l} CO \cdot OCCl_3 \\ | \qquad\qquad = 3\ COCl_2 + CO \\ CO \cdot OCCl_3 \end{array}$$
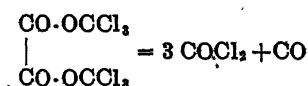
while Henry in 1845 pointed out that phosgene could be produced by the dry distillation of trichloracetic acid, thus:

$$\begin{array}{l} COOH \\ | \qquad\qquad = HCl + CO + COCl_2 \\ C \cdot Cl_3 \end{array}$$
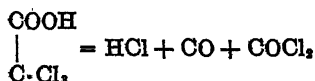
Later, in 1863, Schutzenberger[7] investigated the action of platinum sponge on heated chlorine and carbon monoxide. He

[4] Z. Chem. (2), 4, 5 (1868).
[5] Gilbert's Annalen. 48, 161, (1814).
[6] Ann. Phys. Chem. (3), 19, 352.
[7] C. R. 66, 747.

stated that phosgene was produced, and also that it could be formed either by heating carbon tetrachloride with zinc oxide under pressure, or by passing a mixture of carbon monoxide, chlorine, and carbon tetrachloride vapour through a red hot tube.

$$ZnO + CCl_4 = COCl_2 + ZnCl_2$$
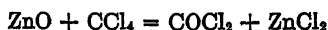
In 1869[8] he obtained phosgene by the interaction of chlorine monoxide and carbon bisulfide:

$$CS_2 + 3Cl_2O = = 2SOCl_2 + COCl_2$$
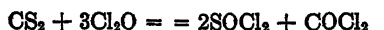
while Emmerling and Lengyel[9] in the same year obtained phosgene, mixed with many other products by passing a mixture of carbon oxysulfide and chlorine through a red hot porcelain tube.

THE PREPARATION OF PHOSGENE

The methods for the preparation of phosgene resolve themselves into the following chief groups:

1. The photochemical combination of carbon monoxide and chlorine.
2. The oxidation of chlorinated hydrocarbons with chromic acid.
3. The interaction of sulfur trioxide or oleum with chlorinated hydrocarbons.
4. The combination of carbon monoxide and chlorine in the presence of a solid catalyst.

The photochemical combination of carbon monoxide and chlorine has been the subject of several detailed investigations, mainly by Chapman and his co-workers, and Weigert. Chapman and Gee[10] experimented on the action of light on mixtures of equal quantities of pure carbon monoxide and chlorine, and found the reaction to be homogeneous, if the surface of the glass be not large. Glass, they found, acted as a weak catalyst, the rate of formation, in the presence of a comparatively large glass surface being 1.237 times as fast as that in an ordinary glass tube. The

---

[8] Ber. 2, 219.
[9] Ber. 2, 546.
[10] J. C. S. 99, 1726.

glass surface was obtained by packing the reaction tube with glass wool. Small quantities of certain substances, such as oxygen, nitric oxide, and especially ozone act as inhibitors, in some cases almost stopping the reaction, although their presence does not, apparently, affect the rate of thermal formation. Weigert[11] investigated the dissociation of phosgene with and without the presence of light, and found that although the presence and wavelength of light alters the rate at which the equilibrium is reached, it had no effect on the position of the equilibrium. This work, in so far as it relates to the wavelength of light in accelerating formation and decomposition of phosgene, was confirmed by the work of Bertholet and Gaudechov[13] whose experiments consisted in placing pure phosgene over dry mercury and exposing the tube to a strong source of ultra-violet light. As decomposition of the phosgene took place the surface of the

TABLE 1

MATERIAL OF TUBE	TIME TAKEN TO PRODUCE FILM ON MERCURY
Clear quartz	5 seconds
Uviol glass	80 seconds
Ordinary glass	More than two hours

mercury was corroded by the chlorine liberated. Tubes of different materials, allowing different amounts of short wave light to pass were used in different experiments, and the time taken for a definite film to form on the mercury was noted. Their results are shown in table 1. It is clear from these results, that ultra-violet light has a strong influence on the decomposition rate of phosgene. Coehn and Becker[12] utilised a "streaming" method for proving the same facts. They passed pure phosgene through a clear quartz tube, at room temperature, illuminating it with the light from a mercury vapour lamp. Dissociation to the extent of 4 per cent was observed in the issuing gas. In another experiment they substituted a tube of "Uviol" glass for

[11] Ann. Physik. (4), **24**, 55 and 243.
[12] Ber. **43**, 130.
[13] C. R. **156**, 1243.

tḥe quàrtz tube. "Uviol" while allowing some of the ultraviolet light to pass, cuts out waves shorter than $256\mu\mu$. With this glass they observed a dissociation of 0.46 to 0.5 per cent.

The oxidation of carbon tetrachloride or chloroform by means of a mixture of potassium bichromate and concentrated sulfuric acid, although too expensive for commercial purposes, furnishes a convenient method for the laboratory preparation of the gas. It was first described by Emmerling and Lengyel[14] in 1869, who warmed chloroform with chromic acid mixture, and obtained phosgene which was passed through a tube containing metallic antimony to remove the excess of chlorine. They assumed the reaction to go:

$$2\ CHCl_3 + 3\ O = 2\ COCl_2 + H_2O + Cl_2$$
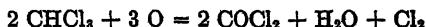
but it was shown by Erdmann[15] and others that the reaction in reality proceeds thus:

$$2 \cdot CHCl_3 + CrO_3 + 2\ O = 2\ COCl_2 + CrO_2Cl_2 + H_2O$$
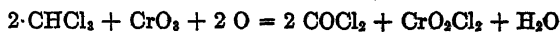
This author describes a convenient laboratory method for the preparation of phosgene from "oleum" and carbon tetrachloride. The apparatus is shown in figure 1. Carbon tetrachloride (120 grams) is placed in a round-bottomed flask, warmed by a waterbath. The flask is connected with an efficient condenser to return any volatilised tetrachloride to the reaction vessel, and "oleum" containing 80 per cent of SO_3 (120 grams) is allowed to drop in slowly from a tap-funnel inserted through the upper end of the condenser tube. The phosgene is washed with concentrated sulfuric acid and condensed in a thick glass vessel, by means of a snow and salt mixture.

This reaction,—between sulfur trioxide, and carbon tetrachloride, was discovered in 1869 by Schützenberger[16] who noticed that when sulfur trioxide and carbon tetrachloride were mixed

[14] Ann. Suppl. 7, 101.
[15] Ber. 26, 1990.
   C. R. 169, 17.
   Gazetta (1920), 5, i, 30.
[16] C. R. 69, 352.

the smell of phosgene became immediately apparent, and a steady stream of the gas was evolved on gently warming the mixture, sulfuryl chloride remaining behind in the residue.

In the same year Dewar and Cranston[17] obtained phosgene by heating together a mixture of chlorsulfonic acid and chloroform:
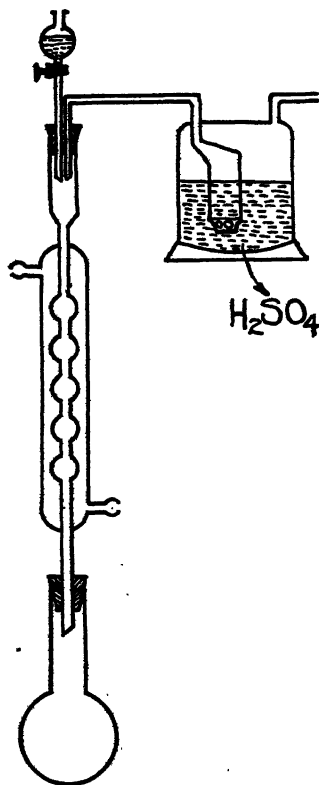
$$CHCl_3 + HSO_3Cl = COCl_2 + 2 HCl + SO_2$$

H_2SO_4

FIG. 1

and the work was repeated by Armstrong[18] who added the fact that hexachlorethane behaved in the same manner:

$$C_2Cl_6 + SO_2 = SO_2Cl_2 + 2 COCl_2$$

[17] Chem. News. **20**, 174.
[18] Ber. **3**, 730.

"Oleum" can take the place of pure sulfur trioxide, the most suitable concentration of SO_3 being 45 per cent (corresponding to the acid $H_2S_2O_7$) when the reaction takes the course:[19]

$$CCl_4 + H_2SO_4 \cdot SO_3 = COCl_2 + 2 SO_3 \cdot HCl$$
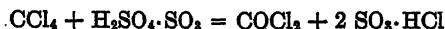
while if less than 45 per cent of sulfur trioxide is present in the "oleum" the excess of sulfuric acid over that required for the above equation remains unaltered. If the sulfur trioxide is in excess of 45 per cent, this excess is said to react according to the equation:

$$CCl_4 + 2 SO_3 = COCl_2 + S_2O_5Cl_2$$
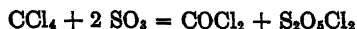
Maugin and Simon[27] expressed the reaction between carbon tetrachloride and sulfuric acid of less than 100 per cent strength by the following equation:

$$(1 + x) CCl_4 + H_2SO_4 + xH_2O = SO_2HCl + (1 + x) COCl_2 + (1 + 2x) HCl$$

This process, using as it does, large quantities of "oleum" and carbon tetrachloride cannot be considered commercially, although its use attained commercial dimensions in Italy, during the late war, when large quantities of phosgene were required for offensive purposes. The reaction between phosphorus pentoxide and carbon tetrachloride was discussed by Gustavson.[20] He found that when one molecule of phosphorus pentoxide reacted with two of carbon tetrachloride phosgene and phosphorus oxychloride were produced, but that when three molecular proportions of the carbon tetrachloride were present no phosgene was observed:

$$\text{i. } P_2O_5 + 2 CCl_4 = COCl_2 + CO_2 + 2 POCl_3$$
$$\text{ii. } 2 P_2O_5 + 3 CCl_4 = 3 CO_2 + 4 POCl_3$$
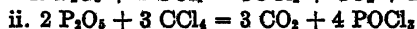
The process which is exclusively used for the large scale manufacture of phosgene is that involving the combination of carbon monoxide and chlorine in the presence of a suitable catalyst. This process was discovered by Paterno[21] who passed the dried

[19] J. Russ. Phys. Chem. Soc. 52, 1.

[20] Z. Chem. (2), 7, 615.

[21] Gazetta. 8, 233.

gases through a tube packed with animal charcoal, and found that their combination was attended by so great an evolution of heat that artificial cooling of the tube had to be resorted to. Atkinson, Heycock and Pope[22] showed that to obtain the best effect the charcoal had to be specially prepared, since many commercial samples of animal and vegetable charcoals were almost catalytically inactive. They found that the best method of preparation was as follows. Freshly crushed ox-bones were heated in sand until no further volatile matter was produced. The residual charcoal was then extracted with hot hydrochloric acid washed with water and reheated in sand. The bone charcoal was further heated in a current of chlorine for several hours. The investigators found that with a U-tube packed with this charcoal and maintained at 40° to 50° combination took place so rapidly that the gases could not be fed in fast enough. The effect of diluting one of the gases with hydrogen was examined, and it was found that if the temperature was kept below 70° no hydrogen chloride was produced; at 80° a slight amount was obtained, and at 90° and above, the formation of hydrogen chloride was considerable.

Among the other reactions, of less importance, that give rise to the formation of phosgene, the following are the most outstanding:

1. The action of carbon monoxide on platinic chloride, to give the substance $Pt \cdot CO \cdot Cl_2$, which decomposes on heating into spongy platinum and phosgene.[23, 24]

2. The decomposition of certain oxides with carbon tetrachloride[25] at temperatures about 300°. Thus:

$$CCl_4 + MO = COCl_2 + M \cdot Cl_2$$

3. Heating the lead ore "phosgenite" which decomposes to some extent according to the reaction:

$$Pb_2Cl_2CO_3 = 2\ PbO + COCl_2$$
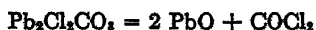
[22] J. C. S. 117, 1410.
[23] Ann. Chem. Phys. (4), 21, 358.
[24] Ann. Suppl. 8, 242.
[25] Z. Angew. Chem. 37, 314.

4. The use of other substances as catalysts in the carbon monoxide-chlorine combination reaction. Thus, Plotnikow[26] found that metal salts, especially anhydrous aluminium chloride would serve instead of charcoal. A mixture of carbon monoxide and chlorine passed over anhydrous aluminium chloride at 30° to 35° gave a good yield of phosgene, which the investigator attributed to the formation and decomposition of a complex $(AlCl_3)_x(COCl_2)_y$.

THE PROPERTIES OF PHOSGENE

Phosgene is a gas without colour, but with a peculiar smell; it does not usually fume in air, although it decomposes water according to the equation:

$$COCl_2 + H_2O = 2\,HCl + CO_2$$

The boiling point of phosgene has been determined by several observers and Beckmann[27] gave it as 8.2°/756 mm. a value which agrees fairly well with that obtained by Atkinson, Heycock and Pope.[22] The vapour pressure of phosgene at various temperatures was determined by Paterno and Mazzuchelli and later over a more extended range by Atkinson, Heycock and Pope (loc. cit.). These latter investigators determined the vapour pressure of phosgene at temperatures from −183° to 100° using different forms of apparatus, one for high, and one for low temperature. The apparatus shown in figure 2, was used for the temperatures from −183° to 10°. The thermojunction sealed through the junction A passes into the liquid phosgene, which is distilled into the vessel by attaching a bulb containing pure phosgene to the tube B and, whilst C is lowered distilling the phosgene into D, which is suitably cooled. The measurements of the variation of the vapour pressure with the temperature can be read off directly on the mercury manometer and the thermojunction. For higher temperatures, where the mercury manometer is of no use, the special apparatus shown in figure 3 is used. The pure phosgene was sealed up in vacuo over mercury,

[26] J. Russ. Phys. Chem. Soc. 48, 457.
[27] Zeit. Anorg. Chem. 55, 370.

whilst dipping below the surface of the mercury was placed a mercury manometer of the closed pattern, the inside of which had been previously silvered.   Thus by maintaining the tube in a bath of water at constant temperature the vapour pressure,
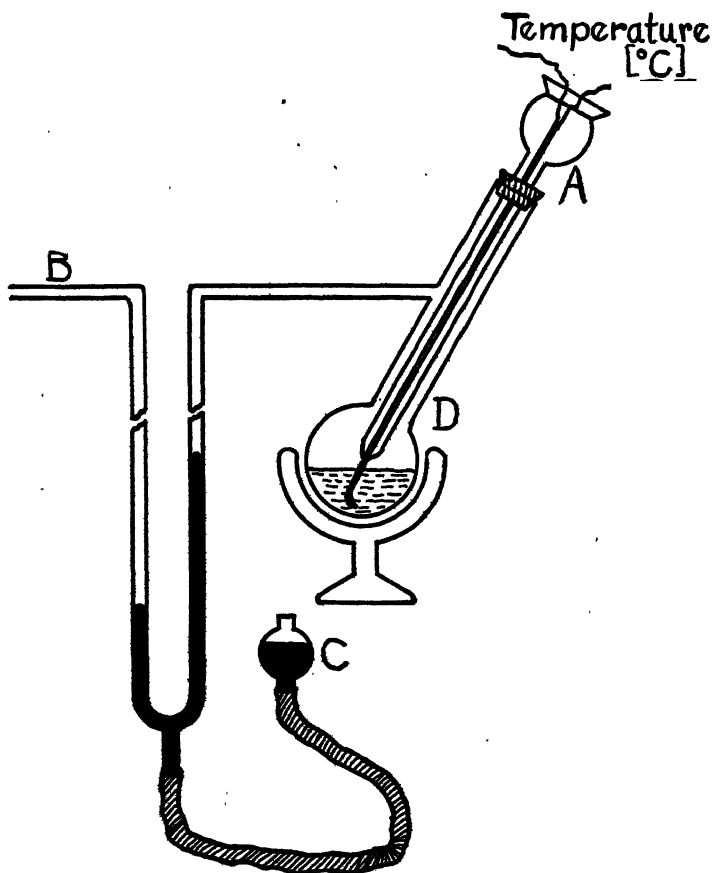

FIG. 2

although not readable directly, could be calculated from the amount of silver removed by the mercury from the inside of the manometer.   The values obtained are given in tabular and graphical form. (See table 2.) The solubility, density and other physical data were very fully investigated by the last

named authors, and their numerical results are given, in part, in table 3. The melting and freezing points of phosgene were found to be −126 and −128° respectively, the figure for the melting point being slightly lower than that [−118°] given by Erdmann.[28] The variation of the density and temperature are given in table 4.
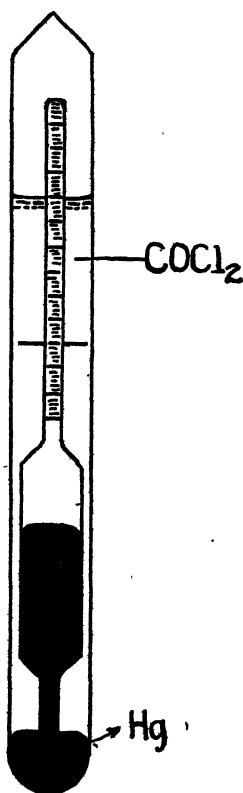
FIG. 3

As a solvent phosgene offers the advantages of dissolving a number of substances which are otherwise difficult to get into solution, and for this reason is a suitable liquid for the determination of molecular weights by the alteration of the boiling point.

[28] Annalen. (1908), **362**, 148.

TABLE 2

METHOD USED	TEMPERATURE	PRESSURE
		mm. Hg
Water-bath.............................	100	16.07 × 760
Water-bath.............................	50	5.11 × 760
Water-bath.............................	17.5	1105.5
Melting ice.............................	0	568.3
Ice and salt mixture.....................	−19	236.0
Freezing point of mercury..............	−39	89.5
Solid CO₂ and ether....................	−79	4.0
Boiling liquid oxygen..................	−183	0.0
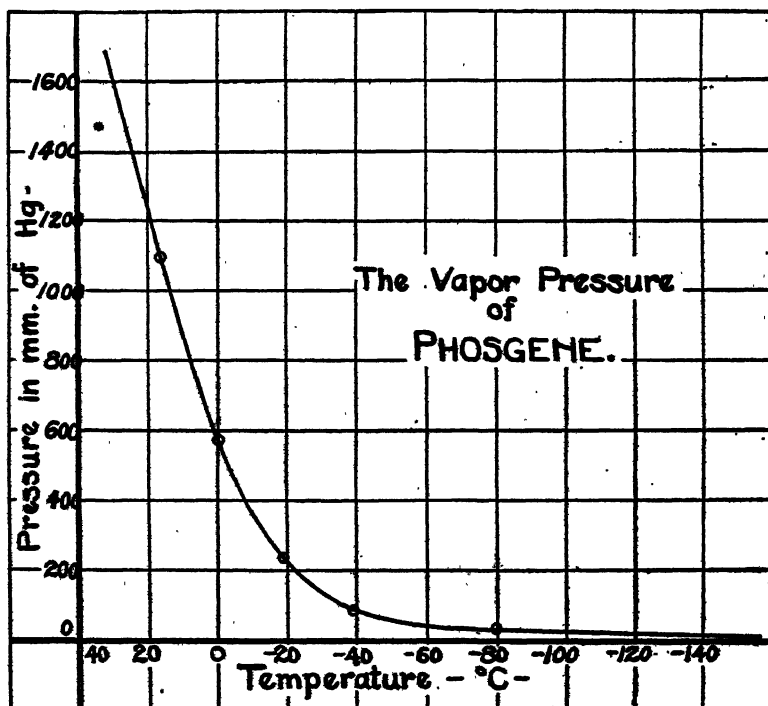
Beckmann and Junker[29] found that the molecular elevation constant was 29, agreeing very well with the value obtained by Atkinson, Heycock and Pope (loc. cit.) by calculation (28–29).

[29] Zeit. Anorg. Chem. **55**, 371.

Beckmann obtained a value of 264 for the molecular weight of iodine by the elevation of the boiling point of phosgene solution indicating the existence of double molecules I_2 in such a solution. In acetic acid he found the value to be 124.

The heat of reaction of phosgene was given by Berthelot[30] as 18,800 from data obtained from the heat of absorption of phosgene in alkali solution. Thomsen[31] pointed out that the absorption of carbon dioxide in Berthelot's experiments was incomplete, and gave as the result of his own work the two values 26,620 and

TABLE 3

Solubilities of phosgene

SOLVENT	TEMP.	VOLS.	TEMP.	VOLS.	TEMP.	VOLS.	TEMP.	VOLS.
Toluene.................	17	244.7	23.5	124.2	30.5	79.39	31.5	79.38
Xylene.................	12.3	457.3	16.4	225.6	16.9	217.9	23.8	103.4
	29.8	71.24						
Creosote oil............	16.2	77.42						
Petroleum b.p. 180–280...	12.3	263.8	15.8	163.1	16.7	143.4	22.4	79.5
Heavy lubricating oil....	23.7	71.2	19.9	49.2	30	48.6		
	15.6	79.7	23.5	39.3	31	24.5		
Nitrobenzene............	16.8	106.4						
a-Chloronaphthalene.....	17.0	104.5						
Chlorobenzene...........	12.3	422.1	16.6	204.1	16.7	221	24.2	99.9
Acetylene tetrachloride...			16.8	149.7	25.1	89.4	29.9	24.9

25,650 whilst that calculated by Ingleson* is 25,500. Thomsen also gives the following values:

$$C + O + Cl_2 = 55140 \text{ cals.}$$
$$CO + Cl_2 = 26140 \text{ cals.}$$
$$COCl_2 + Aq. = 57970 \text{ cals.}$$
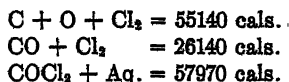
THE DISSOCIATION OF PHOSGENE

The dissociation of phosgene into carbon monoxide and chlorine is extensive at temperatures above 300° and the extent of such dissociation has been determined by several investigators using various means. The earliest work was that of Bodenstein

[30] C. R. (1878), 87, 571.
[31] Ber. 16, 2619, (1883).
* Private communication.

and Dunant[32] who worked at the temperatures of 503°, 553°, 603°, and obtained values which can only be considered as a rough approximation to the truth. In 1909 Horak[33] published a series of results over temperatures from 384° to 500° and obtained a set of results which correspond to a higher degree of dissociation than that which has been found to exist, while in 1920 Atkinson, Heycock and Pope (loc. cit.) published a very irregular set of results which err on the other side, and express a degree of dissociation less than that observed in practice. Their results suffer from the fact that with the method used it was impossible to freeze the equilibrium before partial recombination of the gases had taken place.

TABLE 4

TEMPERATURE	DENSITY PER CUBIC CENTIMETER	TEMPERATURE	DENSITY PER CUBIC CENTIMETER
	grams		grams
−110	1.685	−20	1.481
−100	1.633	−10	1.459
−90	1.640	0	1.435
−80	1.617	10	1.412
−70	1.594	20	1.388
−60	1.572	30	1.363
−50	1.549	40	1.338
−40	1.526	50	1.314
−30	1.504		

Recently, the matter has been cleared up to a great extent by two series of careful and painstaking researches, by Bodenstein and Plaut[34] on the one hand, and by Ingleson on the other. The former investigators obtained results which correspond very well with those of the latter, who by working at carefully regulated temperatures, and taking great precautions that the true equilibrium was the one examined, were able to obtain very consistent results. In a long series of preliminary experiments Ingleson found that trustworthy results could not be obtained

[32] Z. Phys. Chem. 61, 437.
[33] (See 34).
[34] Z. Phys. Chem. 110, 399.

by the heating of phosgene in glass bulbs and that the use of a quartz bulb was necessary. They also found that more reliable results were obtained by the measurement of the actual disso-

TABLE 5

TEMPERATURE	K_c	AUTHOR
603	15.0	BD
553	38.0	BD
506	612.5	AHP
505	379.0	AHP
505	105.4	AHP
503	78.2	BD
500	137.5	H
486	209.5	AHP
481.4	151.6	I
475	223.9	H
460	298.7	AHP
451.3	264.8	BP
450	343.9	H
449	425.0	AHP
444.1	348.2	I
443	419.1	AHP
425	630.7	H
415	717.9	I
413.6	719.5	BP
406	1414.0	AHP
404	1123.0	H
400	954.6	AHP
399	1365.0	H
394.6	1230.0	BP
389	1409.0	I
384	2054.0	H
373.3	1884.0	BP
357	3654.0	I
341	8281.0	AHP

BD, Bodenstein and Dunant, BP, Bodenstein and Plaut; AHP, Atkinson Heycock and Pope; I, Ingleson; H, Horak.

ciation pressure than by the freezing of the equilibrium and estimation iodometrically of the free chlorine.

A table containing the values of K_c, the dissociation constants at various temperatures is given (table 5).

Perkin[35] in his examination of the magnetic rotation of organic compounds, investigated the series, (a) phosgene, (b) ethyl chlorocarbonate and (c) diethyl carbonate and found that the difference between the magnetic rotations of the first pair was unequal to that between the values of the second pair. · Delépine[36] in a series of researches dealing with the variations of physical constants obtained by substituting sulfur for the oxygen of organic compounds was not able to obtain very conclusive results. A typical series of values correlating boiling point and constitution gave the results in table 6. The critical constants of phosgene were obtained by Hakespill and Matthiesen[37] by carefully heating pure phosgene enclosed in a thick glass tube embedded in electrically heated copper and aluminium blocks. Their

TABLE 6

OXYGEN COMPOUND	B.P.	SULPHUR COMPOUND	B.P.	DIFF.
Carbon dioxide	−79	Carbon disulphide	47	2 × 63
Carbonyl chloride	8	Thiocarbonyl chloride	73	65
Acetic anhydride	137	Thioacetic anhydride	157	20
Furane	32	Thiophene	84	52
Phosphorus oxychloride	110	Phosphorus thiochloride	124	14

value for the critical temperature is $183 \pm 0.5°$ comparing well with the value 190°[*]calculated from the Ostwald equation:

$$\alpha = l \, (2T_c - T)$$

One of the peculiar physical properties of phosgene is its ability to arrest certain isomeric changes. In 1908 Lowry and Magson[38] noted that certain nitrocamphor solutions in chloroform failed to give mutarotation phenomena. The presence of small traces of acid in the chloroform was also observed to arrest this change (N/100 acetic or hydrochloric). The active agent was shown to be the phosgene produced in an acid solution of chloroform. The presence of an extremely small trace of piperidine (N/10,000) gave chloroform, solutions in which showed, at first, rapid mutaro-

[35] Chem. News. 69, 224 (1894).
[36] C. R. 153, 727.
[37] Bull. Soc. Chim.
[38] J. C. S. 93, 119, (1908).

tation, but which on keeping, gradually lost the ability of making active solutions. On the addition of further piperidine, however the activity was restored.

<center>REACTIONS OF PHOSGENE WITH INORGANIC COMPOUNDS</center>

<center>*With elements*</center>

Phosgene reacts with a number of metallic elements, especially on heating, to give the chloride of the metal and carbon monoxide. Many of the light elements—sodium, potassium etc.,—react at the ordinary temperatures, but zinc, magnesium etc., only react on warming.[39] It was pointed out by Ribeau[40,41] that this reaction was probably the basis of the older method for the preparation of metallic chlorides by heating a mixture of the oxide and charcoal in a current of chlorine. It is supposed that the oxide and chlorine in the presence of carbon give phosgene, which reacts with the metal or metal oxide to give the chloride.

It has been found, also, that phosgene is capable of exciting the emission of electrons from the surface of sodium and potassium. Thus when sodium or potassium,[42] or an alloy of the two metals, is exposed to a very low pressure of phosgene, electrons are emitted and the metal acquires a positive charge, even in the complete absence of light. The metal is charged to the extent of about one volt. The phenomenon has also been observed by Richardson.[43]

The interaction of carbonyl chloride and fluorine is of interest in that it gives rise to a substance which may possibly be carbonyl difluoride. This reaction is scarcely to be expected in view of the relative strengths of the halogens, and Humiston[44] observed that fluorine had no action on liquid phosgene. When, however, fluorine and phosgene are passed through a copper tube filled with calcium fluoride heated to 200° and the issuing gases con-

[39] Berzelius' Jahresberichte. Fort. Phys. Wiss. **16**, 162 (1837).
[40] C. R. **151**, 1432.
[41] C. R. (1892), 1160.
[42] Ann. Physik. **36**, 308.
[43] Trans. Roy. Soc. **222A**, 1, 43.
[44] J. Phys. Chem. **23**, 572.

densed, a yellow liquid is obtained B.P. −42°. It is explosive and highly reactive, and is thought to be carbonyl fluoride COF_2. Attempts made during the war for the production of carbonyl fluoride on a large scale failed.

With aluminium halides

The reaction of phosgene with aluminium halides gives rise to some interesting compounds. Baud[45] investigated the interaction of phosgene with aluminium chloride (anhydrous), and obtained, in the cold, a solid compound which analysed out at $Al_2Cl_6 \cdot 5COCl_2$. On warming the compound to 30° the vapour pressure becomes equal to 760 mm. and phosgene is evolved until the formula of the remaining compound is $Al_2Cl_6 \cdot 3COCl_2$. This latter compound is a syrupy liquid solidifying at 9° but it loses a further two molecules of phosgene at 55° giving a crystalline compound $Al_2Cl_6 \cdot COCl_2$, which retains the last molecule of phosgene up to 150°. It was the existence of these complexes that led to the hypothesis that similar complexes between aluminium chloride and halogenated compounds were responsible for the Friedel-Crafts reaction. With anhydrous aluminium iodide[46] there is a strong reaction even at the ordinary temperatures, while, when phosgene is passed into aluminium iodide heated to 200° in an air-bath, a thick fluid is obtained crystallising to a brown solid on cooling. It analyses out as $Al_3(CO)_2 \cdot Cl_2 \cdot I$.

CARBONYL BROMIDE

Besson[47] made an unsuccessful attempt to prepare carbonyl bromide by the interaction of phosgene and aluminium bromide. A very small quantity of a yellow liquid B.P. 63° to 66° was obtained. Later Bartal[48] was more successful, and by reacting phosgene with a large excess of aluminium iodide obtained a yellow liquid which on fractionation gave a yellow liquid B.P.
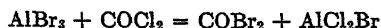
[45] C. R. 140, 1688.
[46] C. Z. Anorg. Chem. 56, 49.
[47] C. R. 120, 190.
[48] Z. Anorg. Chem. 55, 152.

58° to 60° consisting of a solution of bromine in carbonyl bromide. A reddish-brown solid remained behind which had the composition $AlCl_2Br$, so that, presumably, the reaction takes the course:
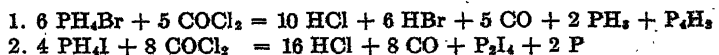
$$AlBr_3 + COCl_2 = COBr_2 + AlCl_2Br$$

If the phosgene be in excess of the aluminium bromide, a reaction takes place, but no carbonyl bromide can be isolated, the compound $COClBr$, being exclusively produced. Carbonyl bromide dissociates very easily at ordinary temperatures and cannot therefore, be obtained free from bromine. In this connection it may be noted that Berthelot[49] had previously, (and unsuccessfully) tried to prepare carbonyl bromide by the photochemical combination of bromine and carbon monoxide. It was found by Bonhoeffer[50] that carbonyl bromide was decomposed by light, even when the latter had been filtered through bromine vapour; the rapidity of the decomposition was almost too great for measurement, and violated Einstein's Law. It may be of interest to recapitulate the work done on carbonyl bromide. In 1905 Bartal[51] prepared the substance in some quantity by reacting phosgene and boron tribromide at −20°, when the reaction proceeded according to the equation:

$$BBr_3 + COCl_2 = COBr_2 + COClBr + BCl_3$$

The mixture was distilled and the following fractions collected:

1.	−20° to 12°	Excess of phosgene
2.	12° to 20°	Boron trichloride
3.	30° to 40°	Carbonyl chlorobromide
4.	60° to 70°	Carbonyl bromide

Besson[52] obtained a very small quantity of carbonyl bromide by the passage of a mixture of phosgene and hydrogen bromide through a tube heated to 200°. Other attempts, such as the interaction of phosgene and phosphonium bromide were unsuccessful. In the latter case the reaction proceeded thus:
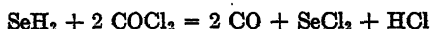
1. $6 PH_4Br + 5 COCl_2 = 10 HCl + 6 HBr + 5 CO + 2 PH_3 + P_4H_2$
2. $4 PH_4I + 8 COCl_2 = 16 HCl + 8 CO + P_2I_4 + 2 P$

[49] C. R. 87, 571.   [51] Ann. 345, 334.
[50] Z. Physik. 13, 94.   [52] C. R. 120, 140.

An attempt to prepare selenophosgene by the reaction:

$$SeH_2 + COCl_2 = CSeCl_2 + H_2O$$

resulted in the reaction:

$$SeH_2 + 2\ COCl_2 = 2\ CO + SeCl_2 + HCl$$

An attempt to[53] prepare carbonyl dicyanide by the action of phosgene upon an aqueous solution of the potassium cyanide at $-10°$ led to the formation of an addition product between potassium cyanide and phosgene to which the formula 1 has been arbitrarily assigned. It
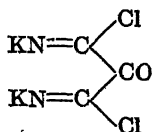
Formula 1

breaks up into potassium chloride, potassium carbonate and azulmic acid. It was also ascertained that phosgene and silver cyanide were mutually inactive at temperatures up to 150°. The action of phosgene on hydrogen peroxide has been shown by Kleinstuck[54] to lead to the formation of formaldehyde according to the equation:
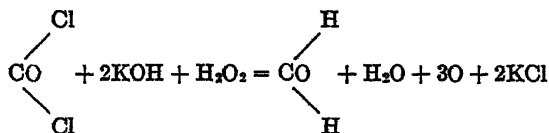
indicating that it behaves as the true acid chloride of carbonic acid.

PHOSGENE AND OXIDES

By heating metal oxides and charcoal in a current of chlorine, it was found possible to prepare the chlorides,[55] but a simpler

[53] Nef. Ann. 287, 309.
[54] Kleinstueck. B. 51, 108.
[55] Œrstedt. Ober. D. vid. Serbsks. Fort. 25, 1824.

and neater way is to heat the oxide in a current of phosgene when the pure anhydrous chloride is obtained.[56] Examples of this reaction with the optimum temperature are given in table 7. Chauvenet[57] points out that in many cases the chloride sublimes in a very pure state; thus, for example, thorium tetrachloride sublimes in fine prismatic needles, stannic chloride distils over and ferric chloride sublimes in deep garnet red crystals.

Sulfides react equally easily with phosgene[58] giving the chloride of the metal and carbon oxysulfide according to the general equation:

$$MS + COCl_2 = COS + MCl_2$$
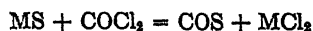
TABLE 7

OXIDE	TEM-PERA-TURE	CHLORIDE OBTAINED	OXIDE	TEM-PERA-TURE	CHLORIDE OBTAINED
Vanadic acid	350	$VaCl_4$	Beryllium	450	$BeCl_2$
Tungstic acid	250	WO_2Cl_2	Aluminium	400	$AlCl_3$
Tantalic acid	400	$TaCl_5$	Iron	350	$FeCl_3$
Titanic acid	400	$TiCl_4$—TiO_2Cl_2	Chromium	600	$CrCl_3$
Silicon dioxide		No reaction	Manganese	450	$MnCl_2$
Zirconia	400	$ZrCl_4$	Nickel	550	$NiCl_2$
Thoria	650	$ThCl_4$	Uranium	450	UCl_4
Tin oxide	400	$SnCl_4$	Cerium	600	$CeCl_4$
Barium oxide	500	$BaCl_2$	Yttrium	600	YCl_3
Magnesium	450	$MgCl_2$	Lanthanum	600	$LaCl_3$
Zinc	450	$ZnCl_2$			

The temperature and products of reaction of the metallic sulfides with phosgene have been recorded (see table 8). This reaction was used by Nuricsau[59] for obtaining a fairly pure carbonyl sulfide. Cadmium sulfide was found the most suitable sulfide and was placed (alone, or mixed with asbestos) in the tube A of the apparatus shown in figure 4. The tube is heated to about 400° and a slow stream of phosgene dried over sulfuric acid is passed when a steady stream of carbonyl sulfide is obtained. This is

[56] Chauvenet. C. R. 152, 87.
[57] Chauvenet. C. R. 147, 1046.
[58] Chauvenet. C. R. 152, 1250.
[59] Nuricsau. Ber. 24, 2967 (1891).

washed with strong caustic potash solution and dried over solid potash. The gas so obtained is about 96 per cent pure.

The natural silicates and phosphates are not so easily decomposed by phosgene as the oxides or sulfides, but on heating to 1000° decomposition takes place with many minerals according to the equation:[60]

$$3 \cdot MO \cdot P_2O_5 + 6\ COCl_2 = 3 \cdot POCl_3 + 6\ CO_2 + 3 \cdot MCl_2$$
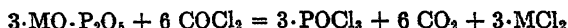
Examples of such decomposition are:

Vivianite. $Fe_3 \cdot (PO_4)_2 \cdot 8H_2O$. . Ferric chloride sublimes.

TABLE 8

| SULPHIDE	FORMULA	TEMPERATURE	PRODUCT
Antimony sulphide	Sb_2S_3	300	$SbCl_3$
Bismuth sulphide	Bi_2S_2	350	$BiCl_3$
Barium sulphide	BaS	400	$BaCl_2$
Zinc sulphide	ZnS	400	$ZnCl_2$
Cadmium sulphide	CdS	400	$CdCl_2$
Copper sulphide	CuS	450	$CuCl_2'$
Mercuric sulphide	HgS	350	$HgCl_2$
Lead sulphide	PbS	350	$PbCl_2$
Iron sulphide	FeS	350	$FeCl_3$
Manganese sulphide	MnS	450	$MnCl_2$
Nickel sulphide	NiS	450	$NiCl_2$

Pyromorphite. $3 \cdot Pb_3P_2O_8 \cdot PbCl_2$ gave lead chloride, $PbCl_2$ some of which sublimed.

Uranite.   Gave volatile uranium tetrachloride, calcium chloride remaining behind:

$$P_2O_5 \cdot 2 \cdot UO_2 \cdot CaO + 8 \cdot COCl_2 = 3 \cdot POCl_3 + 8 \cdot CO_2 + 2 \cdot UCl_4 + CaCl_2$$
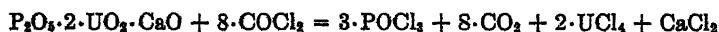
Monazite.   Thorium chloride sublimes.

Silicates react according to the following equation:

$$MSiO_3 + COCl_2 = SiO_2 + CO_2 + MCl_2$$
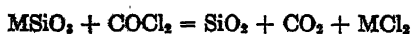
Examples are given in table 9.

Thus the action of phosgene on these minerals offers a method

[60] Barlot and Chauvenet. C. R. **157**, 1153.
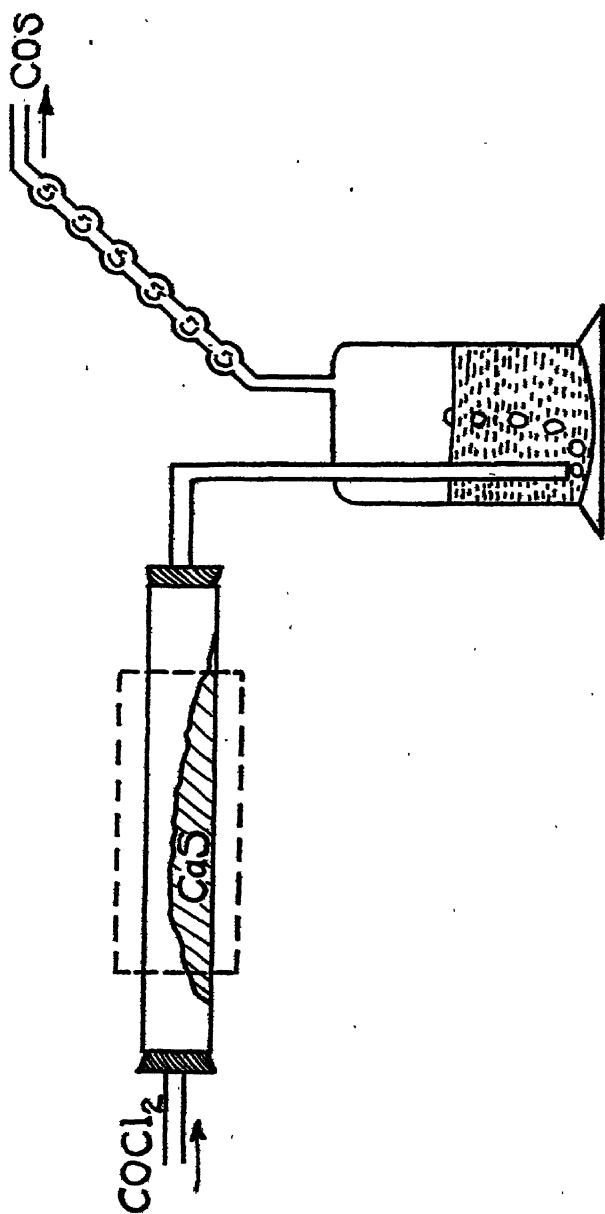
COS

COCl$_2$

CaS

Fig.

for the extraction of thorium from monazite sand and for the "opening up" and analysis of refractory minerals, while it has been proposed to manufacture phosphorus oxychloride by the action of phosgene on precipitated ferric phosphate at 300° to 350°.[61]

THE DETECTION AND ESTIMATION OF PHOSGENE

Phosgene is easily detected qualitatively by aspirating the gas suspected of containing it through a saturated aqueous solution of aniline. After two hours standing the diphenylurea is filtered off and identified by its melting point 99°.[62] If the precipitate is washed with water and dried at 70° the weight gives a comparatively accurate idea of the amount of phosgene present. A variation of this method for quantitative estimation, which gives

TABLE 9

MINERAL	TEMPERATURE	NATURE OF REACTION
Thorite	1000–1150	Chloride sublimes
Gadolinite	1000–1150	Chloride sublimes
Cerite	1000–1150	Chloride sublimes
Zircon	1250	Chloride sublimes
Emerald	1500	No action

more accurate results, consists in treating the precipitate of s-diphenylurea by the Kjeldhal process, estimating the amount of ammonia formed colorimetrically by the means of Nessler's solution.[63]

Chloroform which is to be used surgically for the production of anesthesia, is tested for the presence of phosgene by the aniline test above, which is stated to detect small traces. The corresponding test with p-phenetidine[64] is, however, more sensitive. The chloroform (5 cc.) is dissolved in pure dry benzene (15 cc.) and one drop of p-phenetidine added. A turbidity due to the

[61] Jacobs. U. S. Patent. 1462753.

[62] Kling and Schnitz. C. R. 168, 773.

[63] Kling and Schnitz. C. R. 168, 891.

[64] Scholbein. B. Deut. Pharm. Gesell. 3, 213.

presence of s-di-(4-ethoxy phenyl) urea indicates the presence of phosgene.

The best method of estimating phosgene in a mixture of gases[65] is by absorption. Atmospheric moisture decomposes phosgene comparatively slowly, and acid solutions retard absorption. The method which has been found best is to pass a known volume of the gas (from 2 to 3 litres) in 8 to 10 hours through 10 cc. of 10 N caustic soda solution dissolved in 50 cc. of 95 per cent alcohol. After the gas has been passed, the solution is evaporated on the water-bath and the sodium chloride formed estimated in the usual way. The accuracy of the method depends on the absence of hydrogen chloride or chlorine in the gas. If the phosgene is contaminated with these gases the analysis is correspondingly more difficult. Bertholet had to resort to the following cumbrous series of operations for the estimation of phosgene in admixture with air, carbon monoxide and chlorine. The chlorine was removed by shaking with mercury, the phosgene by warming with potassium, the oxygen by pyrogallate and the carbon monoxide with cuprous chloride. The residue was read off as nitrogen. Various iodometric methods have been used for the estimation of phosgene in the presence of chlorine, for the purposes of determining the amount of dissociation of the phosgene (q.v.). The iodometric methods are not accurate and depend for their action on the assumption that phosgene in passing through a solution of potassium iodide liberates no iodine—an assumption which is scarcely justified by facts.

The presence of chlorine is a source of secondary impurities in phosgene which is packed or transported in iron containers, since it converts the iron into ferric chloride which is soluble in phosgene to the extent of one part in a thousand. Phosgene also occurs as an impurity in commercial samples of titanium tetrachloride. It can be estimated therein by dissolving 5 or 10 cc. of the sample in 125 cc. of dilute hydrochloric acid, and aspirating the carbon dioxide formed by the decomposition of phosgene, through a mixture of 25 cc. of 2 N caustic soda and 50 cc. of N/5 baryta.

[65] Bertholet. Bull. Soc. Chim. N.S. 13, 9, (1870).

The carbonate produced is estimated by the usual titration with acid.

The presence of hydrochloric acid in phosgene is, however, more easily and accurately determined. The following method is used. Finely powdered mercuric cyanide (5 grams) is placed on the bottom of a 1-litre flask furnished with two exit tubes and taps. A sealed bulb containing a weighed amount of phosgene (about 1 gram) is also placed on the floor of the flask and the latter evacuated. The flask is then shaken to break the glass bulb and liberate the phosgene. Any hydrogen chloride present liberates hydrocyanic acid from the mercuric cyanide which is, however, unattacked by phosgene. After 12 to 14 hours the gases are aspirated off through N/2 caustic soda solution, and the cyanide titrated against N/10 silver nitrate solution after the addition of 5 cc. of ammonia solution (d = 0.880) and 1 cc. of 10 per cent potassium iodide solution.

THE REACTIONS OF PHOSGENE WITH AMMONIA AND PRIMARY AMINES

Davy (loc. cit.) in his original paper noted that ammonia and phosgene reacted, but Regnault[66] was the first to investigate the nature of the reaction. He prepared his phosgene by the older photochemical method and observed that four volumes of ammonia were required to neutralise one of phosgene. He isolated from the white substance produced—formerly thought to be a compound of ammonia and phosgene, or "chlorocarbonic acid" as it was then termed—ammonium chloride and a white substance which was at first thought to be urea. It failed, however, to give the precipitate with nitric acid, characteristic of urea, and gave instead a brisk effervescence. These facts led Regnault to suggest that the substance was an isomer of urea to which he gave the provisional name of "carbamide." It seems probable in view of the properties of urea, that the nitric acid used by Regnault was highly contaminated with nitrous acid. Hofmann[67] in 1849 repeated the reaction and obtained ammonium chloride

[66] Regnault. Ann. Chim. Phys. (2). 69, 189 (1838).

[67] Hofmann. Ann. 70, 139 (1849).

and urea as did Natanson[68] seven years later. Both these latter investigators were able to obtain the characteristic precipitate of urea nitrate. Natanson considered this synthesis of urea to be a sequel to the work of Wöhler in 1828 and remarks: "Es entsteht hier also der Harnstoff aus drei sogenannten anorganischen Gasen, das Chlor, Kohlenoxyd und Ammoniak." Bouchardat[69] extended this work on the reaction of ammonia and phosgene and observed the formation of guanidine, cyanuric acid and a compound which he termed "melanuric acid" and which is, in all probability, cyammelide. Fenton[70] proved the presence of urea in the reaction product by the reactions with hypochlorite, by analysis and by the crystalline form. He also identified microscopically some guanidine sulfate.

Gattermann and Schmidt[71] passed phosgene over ammonium chloride heated to 400° in an air-bath. A colourless liquid distilled over which on standing gave bright needles of urea chloride (or carbamic chloride), $NH_2 \cdot CO \cdot Cl$, melting at 50°. The substance has an unpleasant smell, fumes in air and commences to distil at 60° with much decomposition into hydrogen chloride and isocyanic acid. With water the compound decomposes thus:

$$Cl \cdot CO \cdot NH_2 + H_2O = CO_2 + NH_4Cl$$

whilst on standing it decomposes into hydrogen chloride and cyammelide. Its use as a synthetic agent will be discussed later.

Werner[72] has suggested a complicated mechanism for the reaction of phosgene with ammonia. He investigated the yields of various substances when ammonia and phosgene are allowed to react at various temperatures, pure dry ammonia being passed through a 3 per cent solution of phosgene in pure benzene. His results are given in table 10. He argues from this table that since the amount of urea increases as the temperature increases,
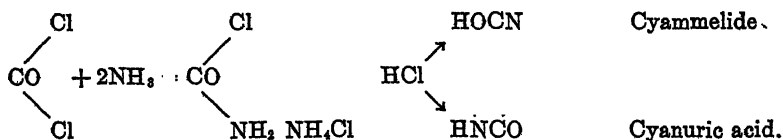
[68] Natanson. Ann. 98, 287 (1856).
[69] Bouchardat. C. R. 69, 962 (1869).
[70] Fenton. J. C. S. 35, 793 (1879).
[71] Gattermann and Schmidt. Ber. 20, 858.
[72] Werner.

it must be produced by rearrangement of ammonium cyanate produced in a somewhat roundabout way:

$$\begin{array}{ccccc}
\text{Cl} & \text{Cl} & \text{HOCN} & & \text{Cyammelide}\,. \\
\diagup & \diagup\,\nearrow & \nearrow & & \\
\text{CO} + 2\text{NH}_3 \cdot \text{CO} & \text{HCl} & & \\
\diagdown & \diagdown\,\searrow & \searrow & & \\
\text{Cl} & \text{NH}_2\ \text{NH}_4\text{Cl} & \text{HNCO} & & \text{Cyanuric acid.}
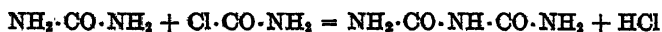\end{array}$$

In view of the extreme local heating obtained in reacting these substances, it seems that the reaction may go partly in the direction indicated, but it is more likely that the urea chloride (if formed) would react immediately with ammonia thus:
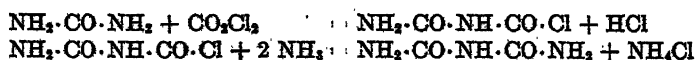
$$2\ \text{NH}_3 + \text{Cl}\cdot\text{CO}\cdot\text{NH}_2 = \text{NH}_2\cdot\text{CO}\cdot\text{NH}_2$$

TABLE 10

	TEMPERATURE		
	20.25	40.5	65.7
Urea	31.7	37.3	41.2
Biuret	14.4	10.1	7.8
Ammelide	7.65	8.6	10.6
Cyanuric acid	3.48	6.4	5.95
Cyammelide	0.69	Trace	None

than that decomposition would take place, followed by urea synthesis from ammonium cyanate formed as fourth product in the chain of reactions. Two other points in Werner's theory of the mechanism of this reaction seem open to criticism. In the first place he accounts for the formation of biuret by the action of cyanic acid on urea, a reaction which does not take place rapidly or quantitatively under any circumstances. It seems far more likely that the biuret is obtained by the action of urea chloride on urea:

$$\text{NH}_2\cdot\text{CO}\cdot\text{NH}_2 + \text{Cl}\cdot\text{CO}\cdot\text{NH}_2 = \text{NH}_2\cdot\text{CO}\cdot\text{NH}\cdot\text{CO}\cdot\text{NH}_2 + \text{HCl}$$

or by the action of phosgene, followed by ammonia on urea

$$\text{NH}_2\cdot\text{CO}\cdot\text{NH}_2 + \text{CO}_2\text{Cl}_2 = \text{NH}_2\cdot\text{CO}\cdot\text{NH}\cdot\text{CO}\cdot\text{Cl} + \text{HCl}$$
$$\text{NH}_2\cdot\text{CO}\cdot\text{NH}\cdot\text{CO}\cdot\text{Cl} + 2\ \text{NH}_3 = \text{NH}_2\cdot\text{CO}\cdot\text{NH}\cdot\text{CO}\cdot\text{NH}_2 + \text{NH}_4\text{Cl}$$

These side reactions would probably be less extensive as the temperature rose, since the compounds containing the $-CO \cdot Cl$ group would probably be very unstable at higher temperatures. This would account for the fact that the combined amounts of urea and biuret is almost independent of the temperature (46.1, 47.4, and 49 in table).

The second point which it is difficult to see is the mechanism of the formation of allophanic ester on shaking the benzene solution with alcohol. Werner postulates the following reaction:

$$2 NH_2 \cdot CO \cdot Cl + C_2H_5 \cdot OH = NH_2CO \cdot NH \cdot CO \cdot O \cdot C_2H_5 + 2 HCl$$

It is more likely to be formed by the esterification of allophanyl chloride (produced as above) thus:

$$NH_2 \cdot CO \cdot NH \cdot CO \cdot Cl + C_2H_5OH = NH_2 \cdot CO \cdot NH \cdot COO \cdot C_2H_5 + HCl$$

PRIMARY AMINES

The primary amines react with phosgene violently with the liberation of much heat, which if not controlled by conducting the reaction in a diluting solvent leads to decomposition of the products. If the amine alone is used, the product depends considerably on the conditions of the experiment. By passing the vapour of phosgene through a solution of the amine in an inert solvent, e.g., benzene, the symmetrical urea is almost always formed:

$$2 \cdot R \cdot NH_2 + CO \cdot Cl_2 = 4 \cdot NH \cdot CO \cdot NH \cdot R$$
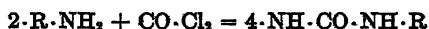
but by spraying an emulsion of amine and water into rapidly stirred liquid phosgene a quantity of the isocyanate can be obtained:

$$R \cdot NH_2 + CO \cdot Cl_2 = R \cdot NCO + 2 \cdot HCl$$

This method of obtaining the isocyanate is not convenient, and the yield is not good. It is much better to heat the well-dried, and powdered amine hydrochloride in a stream of phosgene when the carbamyl chloride is formed which decomposes on heating into hydrogen chloride and the isocyanate:

$$R \cdot NH_2 + COCl_2 = R \cdot NH \cdot CO \cdot Cl + HCl$$
$$R \cdot NH \cdot CO \cdot Cl = R \cdot NCO + HCl$$
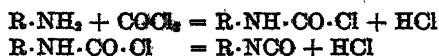
Hofmann[73] was probably the first to observe the action of phosgene on aniline, and records that the carbanilide and aniline hydrochloride were produced. The method, obviously, gives us a way of synthesising almost any symmetrical urea of the type $R \cdot NH \cdot CO \cdot NH \cdot R$. Among many ureas that have been made in this way the following may be mentioned:

s-Di-4-methylphenyl urea   from p-toluidine[74]................m.p. 256
s-Di-4-methoxyphenyl urea from p-phenetidine[75]..............m.p. 174
s-Di-m-cymyl urea   from m-cymidine[76]................m.p. 221
s-Di-4-propylphenyl urea   from p-propylaniline[77]...........m.p. 205
s-Di-4-butylphenyl urea   from p-butylaniline[78].............m.p. 283
s-Di-2-anthraquinonyl urea from α-minoanthraquinone[79]......m.p. 300

The last of these compounds is a yellow vat dye. The first mention of the use of phosgene for converting the amine hydrochloride to the isocyanate is a note by Hentschel[80] to the effect that by heating carbanilide (the intermediate product) with phosgene, phenyl isocyanate was formed and distilled over as an oil. The method was patented by Hofmann and Schoensack.[81] The preparation of β-anthraquinonyl isocyanate, from which orange and red vat dyestuffs can be prepared, has also been protected[82] while the reaction for obtaining the intermediate carbamyl chlorides has been extensively used in synthetic organic chemistry. The value of these compounds lies in their intense reactivity, by the aid of which the following classes of compounds can be synthesised:

1. Urethanes. By the treatment of urea chlorides with alcohols. Thus ethyl urea chloride gives ethyl urethane on treatment with alcohol:

$$C_2H_5 \cdot NH \cdot CO \cdot Cl + C_2H_5 \cdot OH = C_2H_5 \cdot NH \cdot CO \cdot O \cdot C_2H_5 + HCl$$
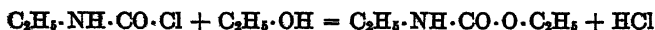
[73] Hofmann. Ann. 70, 139 (1849).
[74] Girard. Ber. 6, 444 (1873).
[75] Muelhaeuser. Ber. 13, 922 (1880).
[76] Kolbe and Warth. Ann. 221, 172 (1883).
[77] Francksen. Ber. 17, 1240.
[78] Pahl. Ber. 17, 1240.
[79] Meister, Lucius and Brüning. G. P. 232739.
[80] Hentschel. Ber. 17, 1284.
[81] Hofmann and Schoetensack. G. P. 29929.
[82] Meister, Lucius and Brüning. G. P. 232739.

Gattermann and Schmidt[23] prepared a whole series of these compounds up to and including the cetyl derivative, but found that in the case of these latter derivatives the reaction proceeded completely to the double compounds thus:
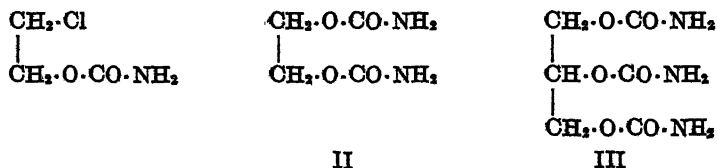
$$Cl + HO \cdot R \qquad\qquad O \cdot R$$

$$CO \qquad\qquad CO \qquad\qquad + 2HCl$$

$$NH_2 + Cl \cdot CO \cdot NH_2 \qquad NH \cdot CO \cdot NH_2$$
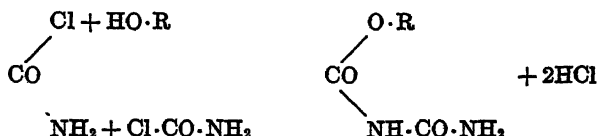
These allophanic esters are produced with the short chain alcohols to some extent. These authors also investigated the action of

TABLE 11

SUBSTANCE	COMPOUND OBTAINED	M.P.
Thiophenol	Phenyl thioallophanate	
a-Naphthol	a-Naphthyl urethane	158
b-Naphthol	b-Naphthyl urethane	187
Thymol	Thymol allophanate	190
Guaiacol	Guaiacyl urethane	127
Pyrocatechol	Pyrocatechyl diurethane	178
Resorcin	Resorcinyl diurethane	194
Hydroquinone	Hydroquinone diurethane	236
Pyrogallol	Pyrogallol triurethane	178
Salicylaldehyde	?	128

urea chloride on ethylene chlorhydrin, glycol and glycerol and obtained the compounds I, II and III, respectively:

$$CH_2 \cdot Cl \qquad\qquad CH_2 \cdot O \cdot CO \cdot NH_2 \qquad\qquad CH_2 \cdot O \cdot CO \cdot NH_2$$

$$CH_2 \cdot O \cdot CO \cdot NH_2 \qquad CH_2 \cdot O \cdot CO \cdot NH_2 \qquad CH \cdot O \cdot CO \cdot NH_2$$

$$CH_2 \cdot O \cdot CO \cdot NH_2$$

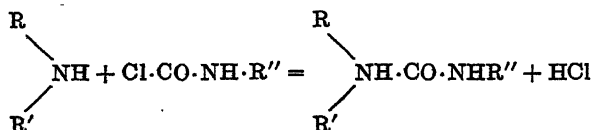$$\qquad\qquad\qquad\qquad II \qquad\qquad\qquad\qquad III$$

Phenols and heterocylic hydroxy compounds also react with urea chlorides to give urethanes. In table 11 is a list of the compounds prepared by these reactions by the last mentioned authors.

[23] Ann. 244, 30.

2. Ureas. This second class of compounds, which may be synthesised by the use of urea chlorides does not need detailed discussion. Any unsymmetrical urea can be obtained by the interaction of a urea chloride with the appropriate amine, thus:

$$\begin{array}{c} R \\ \diagdown \\ \diagup \\ R' \end{array} NH + Cl \cdot CO \cdot NH \cdot R'' = \begin{array}{c} R \\ \diagdown \\ \diagup \\ R' \end{array} NH \cdot CO \cdot NHR'' + HCl$$

3. Aromatic acids. There are several ways in which phosgene can become responsible for the synthesis of aromatic acids and their acid chlorides. Thus the carbamyl chloride can become the acid chloride of the Friedel-Crafts reaction, and lead to the synthesis of the acid through the acid amide. Thus with toluene and carbamyl chloride in the presence of aluminium chloride we get:[34]

$$\langle \bigcirc \rangle CH_3 + Cl \cdot CO \cdot NH_2 = CH_3 \langle \bigcirc \rangle CO \cdot NH_2 + HCl$$
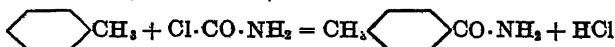
That substitution of the phenyl residue of the aryl amine does not often interfere with the ability of the amino group to react with phosgene is shown by the formation of compounds such as (1),[35] below, which is used as a red-brown vat dye, and (2) which is used as a therapeutic agent in trypanosomiasis, under the name "Fourneau 309":
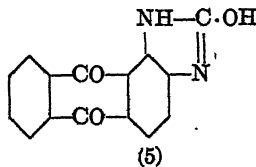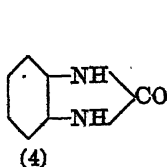

(1)

(2)
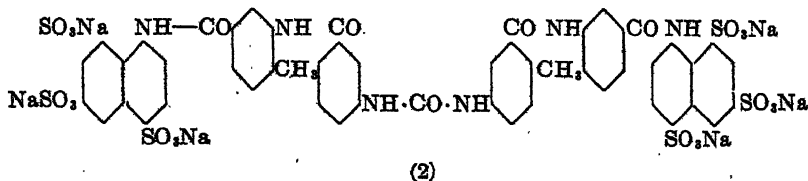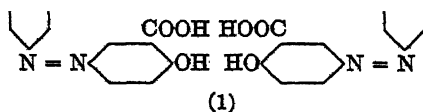
[34] See 83.
[35] B. A. and S. F. G. P. 46737. Fried. 2, 450.

DIAMINES

The reaction between ortho-diamines[86] and phosgene proceeds in exactly the manner expected, giving in the case of ortho-phenylene diamine, phenylene urea a compound occurring in white needles (4)

(4)   (5)

The corresponding 4-tolylene and 3-brom-4-tolylene compounds have also been prepared, while an interesting dyestuff intermediate, stated to have the formula (5) is prepared from 1:2 diaminoanthraquinone.[87] Its constitution as oxy-1:2-anthrimidazole rests on its solubility in alkalies to give orange-red salts.

With p-phenylene diamine[88] it is possible to get the double isocyanate by heating the hydrochloride of the base to 200–250° in a stream of phosgene. 1:4-di-isocyanatobenzene proves to be a white crystalline solid melting at 91°. It gives the usual derivatives on condensation with alcohols and amines, and its vapour density corresponds to the simple formula $C_6H_4 \cdot (NCO)_2$. With benzidin[89] the simple compound with the two isocyanate groups is obtained in a precisely similar manner. It crystallises in splendid needles and gives the usual condensation reactions. m- and o-Tolylene diamines react normally, while compounds of the type (6) react slowly to give cyclic ureas, of which (7) is an example:[90]

(6)   (7)
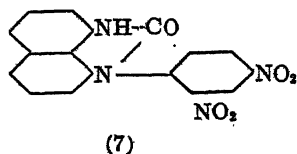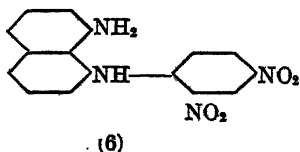
[86] Hartmann. B. **23**, 1046.
[87] Farbenfab. Heyer. G. P. 238981.
[88] Gatterman and Wramplemeyer. Ber. **18**, 2604.
[89] Snape. J. C. S. **49**, 255.
[90] Sachs. and Forster. Ber. **44**, 1744.

SECONDARY AMINES

The action of phosgene on a secondary amine generally leads to the formation of the carbamyl chloride which in the case of the normal secondary amines is a compound of considerable stability. Thus, if phosgene is treated with a well cooled solution of dimethylamine the compound dimethylcarbamyl chloride[91] is obtained as a clear colourless liquid. It is extremely reactive and gives rise to the same condensations as the primary urea chlorides. The use of an excess of the amine leads to the formation of the fully substituted urea. Thus Michler[92] obtained tetraethyl urea from diethylamine and phosgene. He also obtained diphenylcarbamic chloride from diphenylamine and phosgene and found that he could obtain di-, tri- and tetra-

TABLE 12

COMPOUND	SOURCE	M.P.
Methylphenyl carbamyl chloride	Methylaniline[95]	88
Phenyl-b-naphthyl carbamyl chloride	b-Naphthylphenylamine[96]	101-102
Di-β-naphthyl carbamyl chloride	Di-β-naphthylamine[96]	171
Di-benzyl carbamyl chloride	Di-benzylamine[97]	An oil
Di-p-tolyl carbamyl chloride	Di-p-tolylamine[97]	102-103
p-Tolyl benzyl carbamyl chloride	p-Tolyl benzylamine[97]	An oil

phenyl urea by its condensation with suitable amino compounds.[93] The destructive distillation of the as-diphenyl urea prepared in this way proceeds in a peculiar manner giving cyanic acid diphenylamine, diphenylamine cyanide and water:
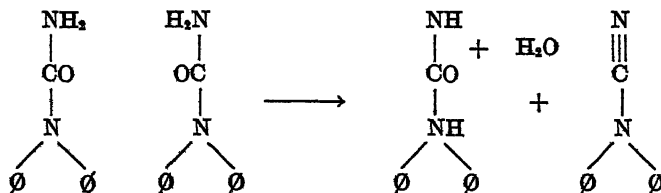
[91] Michler and Escherlich. Ber. 12, 1162 (1879).

[92] Michler. Ber. 8, 1665 (1875).

[93] Michler. Ber. 9, 396 (1876).

This, it may be added, seems to point to the formula $O=C=NH$ for cyanic acid.[94] It has been shown that' mixed aryl-alkyl secondary amines yield similar compounds to the purely aryl compounds, among them being those given in table 12. Among the more interesting experiments with secondary amines are those in which the latter substance is reacted with a compound containing two secondary amino groups in the molecule.

Thus, Michler and Keller[98] found that triphenylguanidine gave the compound (8) and that ethylene diphenyl diamine behaved similarly giving the compound (9), compound (12) being to some extent produced by a side reaction. Hansen[99] obtained similar results with propylene diphenyl diamine (10), obtaining also some of the di-acid chloride (11). The extension of these

(8) (9) (10) (11) (12)
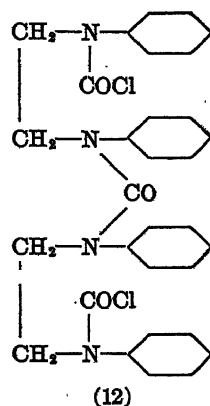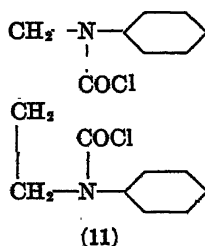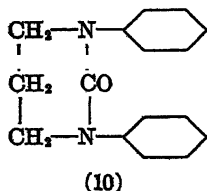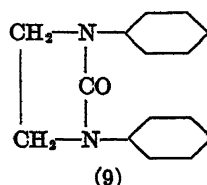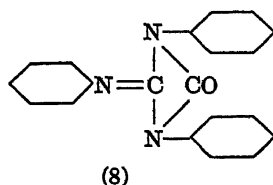
[94] Michler. Ber. 9, 716.
[95] Michler and Zimmermann. B. 12, 1165.
[96] Kym. Ber. 23, 427.
[97] Hammerich. Ph.D. Thesis. Basle. Ber. 25, 1819 (1891).
[98] Michler and Keller. Ber. 14, 2181.
[99] Hanssen. Ber. 20, 781.

experiments to the tetra and penta methylene diphenyl diamines would be interesting, in order to observe whether seven and eight membered rings of this type can be produced. The action of phosgene on the hydrazides is a special case of the series discussed above. Freund and Goldschmidt[100] reacted acetylphenyl hydrazine with phosgene and obtained a compound to which they attached the somewhat improbable formula (13). Other hydrazides (tabulated below) were tried and since the reaction appeared to be a general one the work was extended to the hydrazides of dibasic acids, giving the dicarbazides. Thus in the first investigation the compound from malonyl dihydrazide was accorded one of the formulas (14) or (15).

(13)   (14)   (15)
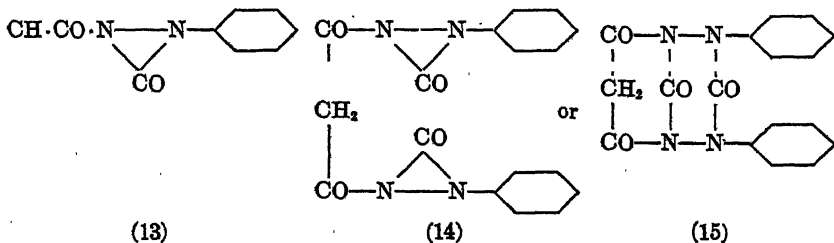
In addition the compounds given in table 13 were prepared.

Freund and Kuh[101] suggested that the formula of the carbazines was better represented by (16) than by the formulae previously used.

(16)   (17)   (18)
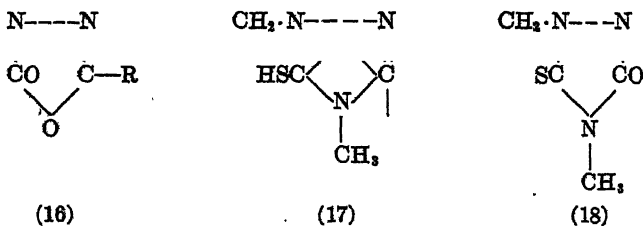
Their formula accorded well with the behaviour of the compounds, and had the merit of avoiding the improbable three-membered ring involved in the earlier formulae. The name "biazolone"

[100] Freund and Goldsmith. Ber. 21, 1240 and 2456.
[101] Freund and Kuh. Ber. 23, 2061.

was given to such structures.   The action, however, of phosgene on the semicarbazides and thiosemicarbazides, as elucidated by Busch and Offermann[102] confirmed the fact that compounds of this type are in reality represented by the cyclic formulae.   In reacting the compound 2-benzyl-4-methyl semicarbazide with phosgene they obtained the compound (21) which they finally identified as 1-benzyl-4-methyl thiazurol. This they found was unstable on heating and passed over into the compound (22) 1-benzyl-4-methyl-5-thiolendoxy triazole. The mechanism of this reaction, which is somewhat obscure was investigated by Busch and Limpach[103,104] who isolated the intermediate com-

TABLE 13

Acetylphenylcarbazine	Acetylphenylhydrazine	
Formylphenylcarbazine	Formylphenylhydrazine	73
Propionylphenylcarbazine	Propionylphenylhydrazine	62
Benzoylphenylcarbazine	Benzoylphenylhydrazine	114
Succinylphenylcarbazine	Succinylphenylhydrazine	225
Malonylphenylcarbazine	Malonylphenylhydrazine	205
Ethylmalonylphenylcarbazine	Ethylmalonylphenylhydrazine	223
Oxalylphenylcarbazine	Oxalylphenylhydrazine	287

pound (20) a thiodiazolone anil, which on fusion or boiling in alcoholic solution changes into the triazole compound thus:

$$\text{R·NHNH}_2 \qquad + \text{CO·Cl}_2 \cdot$$
$$\text{RN}{=}\text{C·SH}$$
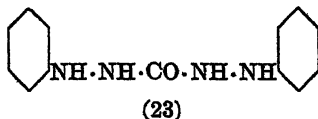
(19)   (20)   (21)   (22)
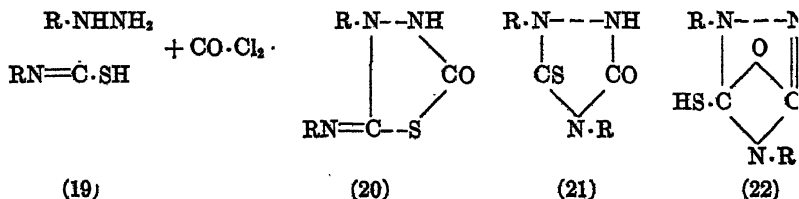
Various of these series of compounds have been prepared, chief among them being those shown in table 14.

[102] Busch and Opfermann. B. 37, 2335 (1904).
[103] Busch. Ber. 42, 4766.
[104] Busch and Limpach. Ber. 44, 569.

The action of phosgene on simple hydrazines does not seem to have been extensively studied. Heller[105] obtained diphenyl-carbazide (23)

(23)

by the action of phosgene on a well cooled solution of phenyl hydrazine. It is a crystalline compound melting at 163°. Acree[106]

TABLE 14

PARENT COMPOUND	THIAZUROL. M.P.	THIOL. M.P.
Benzylmethylthiosemicarbazide	117	157
Benzylallylthiosemicarbazide	108	161
Benzylphenylthiosemicarbazide	147	217
2-m-Tolyl-4-phenylthiosemicarbazide	125	259
2-m-Bromphenyl-4-phenyl thiosemicarbazide	118	257
2-m-Chlorphenyl-4-phenyl thiosemicarbazide	108	260
2-b-Naphthyl-4-phenyl thiosemicarbazide	133	295
2-p-Tolyl-4-phenylthiosemicarbazide	144	240
2-p-Bromophenyl-4-phenyl thiosemicarbazide	170	255

obtained diphenyl carbazyl chloride (24) and tetraphenyl car-bazide by the action of phosgene on as-diphenyl hydrazine.

(23)   (24)
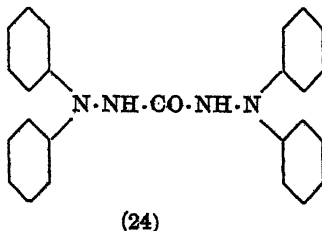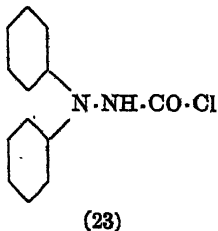
TERTIARY AMINES

Phosgene is without action on the aliphatic tertiary amines, and aromatic tertiary amines react very reluctantly save in the

[105] Heller. Ph.D. Thesis. Ann. 263, 269.
[106] Acree. Ber. 26, 3154 (1903).

presence of anhydrous aluminium chloride, when the Friedel-
Crafts reaction takes place with the formation of a considerable
amount of acid chloride or ketone. Thus with dimethylaniline
the well known Michler's ketone is obtained, tetramethyldia-
minobenzophenone. At the same time a small amount of the
double ketone (26) is obtained.[107]

$$(CH_3)_2 \cdot N \langle \rangle - CO - \langle \rangle - N \cdot (CH_3)_2$$

(25)

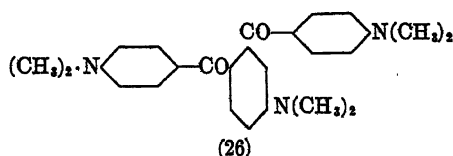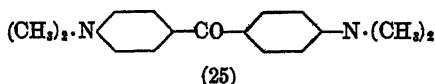$$(CH_3)_2 \cdot N \langle \rangle - CO \begin{array}{c} CO - \langle \rangle N(CH_3)_2 \\ \\ N(CH_3)_2 \end{array}$$

(26)

The former compound—Michler's ketone, can be reduced to the
carbinol (Michler's Hydrol) in the usual way[108] and the two sub-
stances form the starting point of a large series of triphenyl
methane dyes, chiefly blues and violets of the crystal violet type
including Victoria Blue, Night Blue, Wool Green S, Victoria
Blue 4R etc. The ketone can be condensed with the amine
hydrochloride to give the dyestuff, which is said by Hofmann
to be produced by the following series of reactions:[109]

$$(CH_3)_2 \cdot N \langle \rangle - CO - \langle \rangle N \cdot (CH_3)_2 \qquad (CH_3)_2 \cdot N \langle \rangle - C - \langle \rangle N(CH_3)_2$$

$$\underset{}{\xrightarrow{\hspace{2cm}}} \qquad \overset{|}{\underset{Cl \quad Cl}{}}$$

$$(CH_3)_2 \cdot N \langle \rangle - C - \langle \rangle - N \cdot (CH_3)_2$$

$$Cl \cdot \overset{..}{N} \cdot (CH_3)_2$$
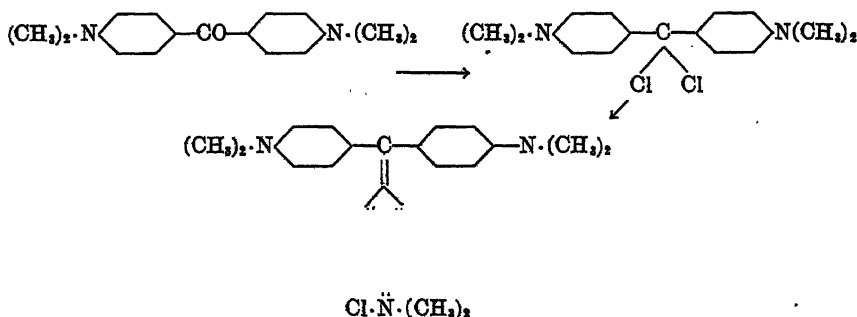
The intermediate ketone need not be isolated, for the tertiary
amine can be directly condensed in excess, with phosgene, in the

[107] Michler. Ber. 9, 716 1899.
[108] See 107.
[109] Hofmann. Ber. 18, 770.

presence of aluminium or zinc chloride to give the dyestuff. Thus in one typical patent[110] on this class of dyestuff, dimethyl-aniline (100 kgm.) is saturated with phosgene (18 to 20 kgm.) at 20° and allowed to stand for 24 hours. Further dimethyl-aniline (50 kgm.) and powdered zinc chloride (30 kgm.) are added. The mixture is then warmed to 40° to 50° and a further 20 kgm. of phosgene passed in, and the whole digested at 60° until a deep paste of methyl violet is formed.[111] The general applicability of such a process is limited by the fact that many substituted tertiary amines do not yield a Michler's ketone. Thus Rassov and Reuter[112] found that on condensing dimethyl-o-toluidine with phosgene a number of compounds including methyl chloride were obtained. The compounds (28) and (29) were obtained but not the compound expected (27).
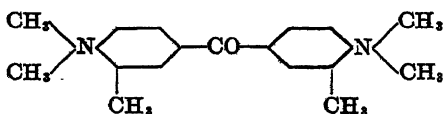
(27)
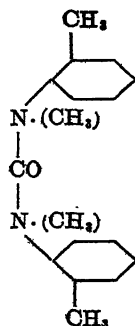
(28)
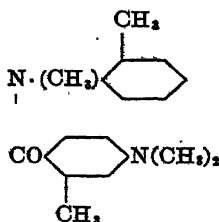
(29)

[110] Meister, Lucius and Brüning. B. P. 8694. J. C. S. I. 7, 205 (1888).
  Badische anilin und Soda-Fabrik. G. P. 26066. Ber. 17, 60.
  Badische anilin und Soda-Fabrik. G. P. 27789. Ber. 17, 339.
  Meister, Lucius and Brüning. G. P. 34463. (1884).
[111] Badische-Anilin und Soda-Fabrik. G. P. 29943. (1884) Ber. 18, 7.
[112] Rassow and Reuter. J. Pr. Ch. 85, 489.

Loeb[113] investigated the action of phosgene on ethenyldiphenyl-tolamine. When one molecular proportion of phosgene and two of the base are used the compound (30) is produced, but with an excess of phosgene the compound (31) is the product of the reaction. On heating it loses a molecule of phosgene giving the cyclic compound (32).
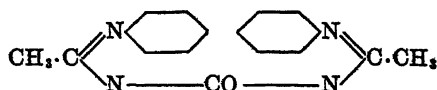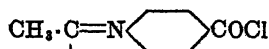
(30)   (31)
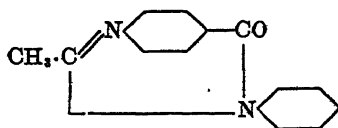
THE REACTION OF ALCOHOLS AND PHENOLS WITH PHOSGENE AND THE SYNTHESIS OF ACIDS AND ACID CHLORIDES

The reaction between alcohols and phosgene is simplicity itself and can be summarized in the equation:

$$R \cdot OH + CO \cdot Cl_2 = R \cdot CO \cdot Cl + HCl$$
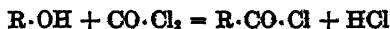
The action, investigated by Dumas[114] has been a standard method for the preparation of chlorocarbonic esters since his time. Dumas extended his observations to methyl chlorocarbonate, and various other investigators have applied the method to other alcohols. Among the chief compounds mentioned are those given in table 15.

By the action of phosgene on glycerol, compounds were iso-lated which included glycerol carbonate (J.C.S., 1925), but the best way of obtaining glycerol carbonate is undoubtedly the action of heat on a mixture of glycerol and phenol carbonate.

---

[113] Loeb. Ber. **18**, 2427.
[114] Dumas. Ann. Chim. Phys. (2), **56**, 226; (2), **58**.

The glycerol carbonate can be distilled out of the mixture in vacuo. It is a white crystalline substance m.p. 148°.[124] The action of phosgene on glycollic ester[125] gives rise to a pulverulent substance which is probably the carbonate. The action of an excess of alcohol on phosgene or of the sodium derivative of

TABLE 15

ALCOHOL	B.P. OF CHLOROCAR- BONATE	REFERENCE
Methyl	71.4	114, 117, 110, 123
Ethyl	90	114, 117, 121
iso-Amyl	154.4	115, 117
Propyl	115.2	116, 117
Butyl		117
iso-Butyl	128.8	118, 117
Ethylene glycol	236	122
Glycolchlorhydrin	158–160	123

TABLE 16

COMPOUND	B.P.
Dimethyl carbonate	90.6
Diethyl carbonate	126
Dipropyl carbonate	168.2
Di-iso-butyl carbonate	190.3
Di-iso-amyl carbonate	228.3
Methyl ethyl carbonate	109.2
Methylpropyl carbonate	130.8
Methylisobutyl carbonate	143.6
Ethylisobutyl carbonate	160.1
Ethylisoamyl carbonate	182.3

[115] Medlock. Quart. J. Chem. Soc. 1, 368 (1849).
[116] Roemer. Ber. 6, 1101 (1873).
[117] Roese. Ann. 205, 229.
[118] Mylius. Ber. 5, 477, (1872).
[119] Klepl. J. Pr. Ch. (2), 26, 448 (1882).
[120] Hentschel. Ber. 18, 1177.
[121] Farb. Bayer. G. P. 118537.
[122] Nemirowsky. J. Pr. Ch. 2, 28, 439.
[123] Nemirowsky. J. Pr. Ch. (2), 31, 173.
[124] Hochstetter. G. P. 252758.
[125] Ann. 154, 257 (1870).

alcohol on that compound leads to the formation of a symmetrically disubstituted ester of carbamic acid:

$$R \cdot O \cdot CO \cdot Cl + NaO \cdot R = R \cdot O \cdot CO \cdot O \cdot R \cdot + NaCl$$
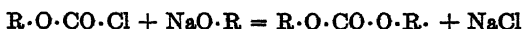
Among the aliphatic carbonic esters obtained by Rosse (loc. cit.) who investigated the whole series very thoroughly, and others, are those given in table 16.

In the aromatic series the action of phosgene on phenols is not nearly so violent as with their aliphatic analogues. Thus phosgene and phenol require to be heated in a sealed tube[126] to obtain phenyl carbamyl chloride $C_6H_5 \cdot O \cdot CO \cdot Cl$. This compound,

TABLE 17

PARENT SUBSTANCE	CARBAMYL CHLORIDE	DI-SUBSTITUTED ESTER OF CARBONIC ADID	REFERENCE
Phenol.................		m.p. 75	[127, 128, 129]
2:4 Dinitrophenol.......		m.p. 125.5	[127]
Cresol (?)...............		m.p. 125	[127]
Eugenol.................	b.p./17 mm. 174	m.p. 93	[130, 131]
Guaiacol................	b.p./10 mm. 110		[130]
b-Naphthol..............	m.p./65–66		[129]
iso-Eugenol.............	b.p./15 mm. 155–157		[130]
Methyl salicylate........	b.p./20 mm. 141–147		[130]
Ethyl-p-oxybenzoate	m.p. 55–56		[130]
Thiophenol..............		m.p. 72	[131]
Salicylaldehyde.........		m.p. 94–95	[131]

however, will condense with ammonia in the usual way, and with amines to give urethanes, and with sodium phenate to give diphenyl carbonate. The symmetrical di-aryl substituted esters of carbonic acid can be most readily obtained by the action of phosgene on the sodium salt of the phenol concerned, preferably in a well-cooled acetone solution. The compounds given in table 17 are among those that have been prepared in this manner.

[125] Kempf. J. Pr. Ch. (2), 1, 402 (1870).
[127] Hentschel. G. P. 24151. Fried. 1, 230, 1880.
[128] Barral and Morrell C. R. 128, 1579.
        Barral and Morel. Bull. Soc. Ch. (7), 21, 725.
[129] Einhorn. G. P. 224108.
[130] Loewenberg. Chem. Zent. 1886, 390.
[131] Hofmann. Z. Ang. Ch. 21, 1896.

With reference to certain of these esters of chlorocarbonic acid Hofmann[131] remarks that they react curiously with pyridine giving rise to compounds of the type (33)

R · O—Ñ—CO—Ñ—Cl
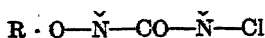
(33)
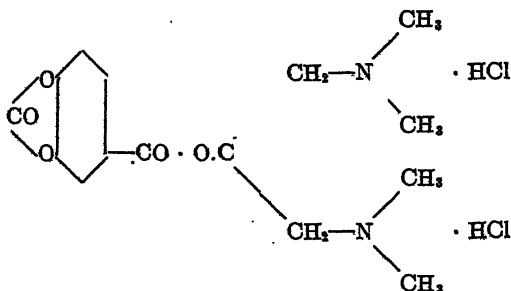
(34)

He also prepared the compound (34) in order to test its therapeutic properties.

It will be noticed that the ortho-dihydric phenols condense
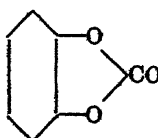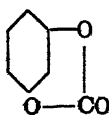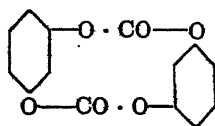
(35)   (36)   (37)

with phosgene to give phenylene carbonic esters as in (35). With the meta- and para- dihydric phenols, the reaction is not so simple. Bernbaum and Lune[132] reported the compound (36) as the product of the reaction, an infusible and insoluble white powder, but it more likely that the compound is in reality the double molecule (37). Many other carbonic esters, e.g., carbonyl gallic ester, etc.,[133] have been prepared, but their mere enumeration serves no useful purpose.

SYNTHESIS OF ACIDS, ACID CHLORIDES AND KETONES

Among the earlier investigators of the reactions of phosgene Harnitz and Harnitzki[134] stated that phosgene and methane

---

[132] Birnbaum and Lune. Ber. 14, 1754 (1881).
[133] Fischer and Freudenberger. Ber. 46, 1116.
[134] Harnitz-Harnitsky. C. R. 60, 923 (1865).

react in sunlight to give a certain amount of acetyl chloride. This was confirmed by Butlerow[135] who observed at the same time that phosgene when reacted with zinc methyl gave acetyl chloride thus:

$$(CH_3)_2 \cdot Zn + 2\ CO \cdot Cl_2 = 2\ CH_3 \cdot CO \cdot Cl + ZnCl_2$$

Bertholet[136] denied that any reaction took place between methane and phosgene at temperatures below a red heat, and his experiments indicated that the same inertness was to be observed in connection with ethane, ethylene, acetylene and benzene, while De Clermont and Fontaine[137] showed a similar lack of reaction between octane and phosgene. The latter investigators seem to be correct, at least for low temperatures and in the absence of catalysts. The process, however, devised by Hochstetter[138] for the preparation of methyl chloride involves the chlorination of methane by phosgene. A mixture of the gases is passed through a tube packed with wood charcoal and heated to 400° when the reaction

$$CH_4 + CO \cdot Cl_2 = CH_3 \cdot Cl + CO + HCl$$

takes place. The methyl chloride is removed from the system by refrigeration.

Phosgene will act as an acid chloride in the Friedel-Crafts reaction, giving the usual products when the reaction is stimulated by the presence of anhydrous aluminium chloride. Harnitz and Harnitzki[139] stated that a small quantity of benzoyl chloride was obtained when phosgene reacted on benzene alone, but this is doubtful; Meyer[140] however, obtained a fair amount of benzoyl chloride by heating silver benzoate and phosgene in a sealed tube. The simplest way of obtaining benzoyl chloride from benzene is to allow a solution of phosgene in the latter

[135] Lehrb. d. Org. Chemie. 297 (1868).
[136] Bertholet. Bull. Soc. Chim. N. S. 13, 9 (1870).
[137] DeClermont and Fontaine. Bull. Soc. Chim. N. S. 13, 494.
[138] Hochstetter. Austr. Pat. Applic. 9887.
    Hochstetter. G. P. 292089. J. Soc. Chem. Ind. 35, 867 (1916).
[139] Harnitz Harnitzky. C. R. 58, 748 (1864).
[140] Meyer. Ann. 156, 271 (1870).

hydrocarbon to stand over anhydrous aluminium chloride.[141] The reaction does not proceed quantitatively to benzoyl chloride, since if the standing is protracted or the temperature rises the second stage of the reaction sets in and benzophenone[142] is formed:

(i) $C_6H_6 + CO \cdot Cl_2$ $HCl + C_6H_5 \cdot COCl$
(ii) $C_6H_5 \cdot CO \cdot Cl + C_6H_6$ $C_6H_5 \cdot CO \cdot C_6H_5 + HCl$
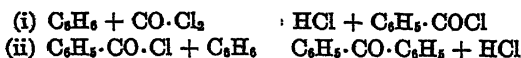
At the same time a small quantity of the compound (38) is formed but its isolation is difficult.

(38)
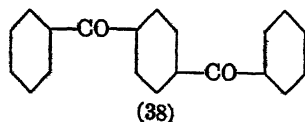
TABLE 18

COMPOUND	REFER-ENCE	M.P. OF ACID FROM THE ACID CHLORIDE	KETONE
Benzene..............................		120	
Toluene..............................	141	177	m.p. 92, b.p. 240
Xylene...............................	144		b.p. 340
Anthracene...........................	145	206	
Anthraquinone........................	145	ca. 300	
Diethylaniline........................	146	188	
Thiophene............................	147		m.p. 87–88.
Methylphenylaminobenzene............	148	184	

The reaction has been recently reinvestigated[143] in an attempt to synthesise anthraquinone by the reaction of an excess of phosgene on benzene or benzoyl chloride. No trace of anthraquinone could be detected.

The reaction between phosgene and aromatic hydrocarbons in

[141] Ador and Krafts. Ber. 10, 2173.
[142] Freidel, Krafts, and Ador. Ber. 10, 1854.
[143] Wilson and Fuller. J. Ind. Chem. Eng. 14, 406 (1922).
[144] Ador and Rilliet. Ber. 11, 399.
    Elbs and Olberg. Ber. 19, 428.
[145] Graebe and Liebermann. Ber. 2, 678 (1869).
[146] Michler and Gradmann. Ber. 9, 1912 (1876).
[147] Gattermann. Ber. 18, 3013.
[148] Sarauw. Ber. 14, 2180 (1881).

the presence of anhydrous aluminium chloride is fairly general and table 18 gives the results obtained with some of the commoner hydrocarbons.

Behla[149] was unable to obtain a carboxylic acid from the interaction of phosgene and dihydroanthracene, and his experiments on the interaction of phosgene and phenanthrenequinone were likewise inconclusive.[150]

Phosgene has been found a convenient agent for the preparation of acid anhydrides and chlorides from the parent acids or their sodium salts. Thus sodium acetate successively undergoes the reactions shown below:

(i)
$$CH_3 \cdot CO \cdot ONa$$
$$+ CO\ Cl_2 =$$
$$CH_3 \cdot CO \cdot ONa$$
$$CH_3 \cdot CO{\Large\diagdown}_{\diagup} O + 2NaCl + CO_2$$
$$CH_3 \cdot CO$$

(ii)
$$CH_3 \cdot CO{\Large\diagdown}_{\diagup} O + CO \cdot Cl_2 =$$
$$CH_3 \cdot CO$$
$$CH_3 \cdot CO \cdot Cl$$
$$+ CO_2$$
$$CH_3 \cdot CO \cdot Cl$$
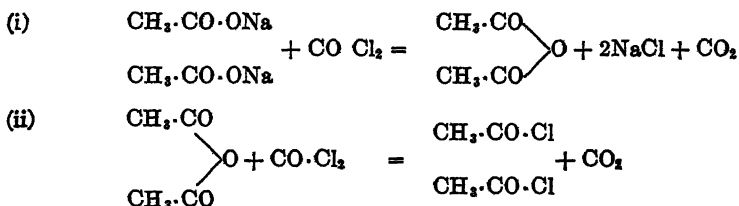
Hofmann and Schoelensack[151] patented this as a general method for the production of acid chlorides and anhydrides of acetic, propionic, butyric and benzoic acids. Later they extended the method to the production of salicylic acid from sodium phenate.[152] The sodium phenate mixed with caustic soda was heated to 140° to 200° in a stream of phosgene. The reaction may be represented:

$$C_6H_5 \cdot ONa\ \ Cl{\Large\diagdown}_{\diagup}CO =$$
$$C_6H_5 \cdot ONa\ \ Cl$$
$$C_6H_5 \cdot O{\Large\diagdown}_{\diagup}CO +$$
$$C_6H_5 \cdot O$$
$$OH$$
$$|$$
$$Na$$
$$=$$
$$C_6H_4 \cdot (OH) \cdot COOH$$
$$+$$
$$C_6H_5 \cdot ONa$$
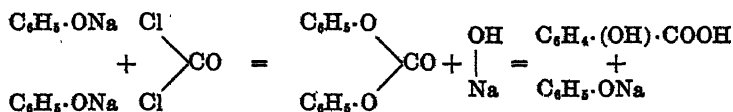
The original general method for the production of the acid chlorides and anhydrides was to heat the sodium salt with phosgene in an autoclave. This method not only gave a poor yield of the desired products but was very difficult of operation

[149] Behla. Ber. 20, 701.
[150] Behla. Ber. 18, 3169 (1885).
[151] Hofmann and Schoetensack. G. P. 29669.
[152] Hofmann and Schoetensack. G. P. 30172.

on a technical scale. To remedy this Hochstetter[153] devised a process in which the vapour of the free acid is passed, together with phosgene through a tube heated to bright redness and filled with a catalyst material (usually wood charcoal). A better yield is obtained while the process is simple and easily controlled.

SOME MISCELLANEOUS REACTIONS OF PHOSGENE

With acid-amides and compounds containing the $-CO \cdot NH_2$ *group*

Among the earlier researches on phosgene was that of Schmidt[154] who investigated its reactions on various compounds of the urea and ureide type. He found that when dry urea was heated under pressure with phosgene that the substance carbonyl di-urea was formed (39):

NH·CO·NH₂   NH·CO·NH   NH·CO·NH·CO·NH₂

CO   CO   CO

NH·CO·NH₂   NH——CO   NH·CO·NH·CO·NH₂
    (39)   (40)   (41)
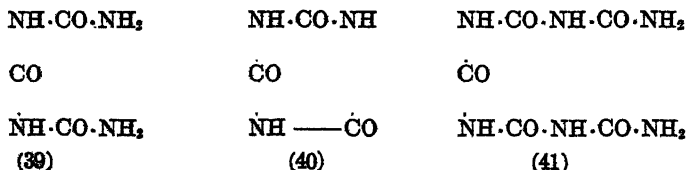
This substance was almost exclusively formed when the mixture was heated at 100° for two days, but on heating carbonyl diurea with phosgene to 150° to 160° cyanuric acid (40) was formed by the loss of one ammonia molecule from the carbonyl-diurea (39).

Biuret behaved in the same way giving in the first instance carbonyl dibiuret (41) which on further heating with phosgene gave cyanuric acid and hydrogen chloride thus:
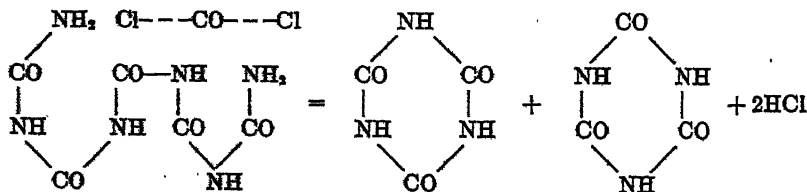
[153] Hochstetter. G. P. 283896. C. Abs. 10, (1916), 93.
      Hochstetter. G. P. 29617. C. Abs. 10, (1916), 94.
[154] Schmidt. J. Pr. Ch. 2, 5, 39 (1872).

Other amides behave in a similar manner, thus acetamide gives diacetylurea (41a) and benzamide gives carbonyl dibenzamide (42).

(41a)   (42)
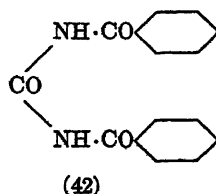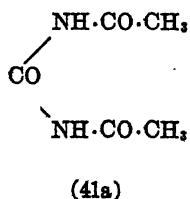
Oxamide was stated to give a mixture of carbonyl diurea, urea and carbon dioxide, but later and more careful investigations by Basarow[155] have shown that when oxamide and phosgene are heated to 170° to 180°C. parabanic acid is formed (43). The earlier work on the reaction between urea and phosgene was confirmed by Schiff.[156] An interesting case, described by Holle-mann[157] is that of the interaction of phosgene and N-dimethyl benzamide, in which there is no free hydrogen on the nitrogen atom. In this case a dichloro compound of the formula (44) was formed.

(43)   (44)
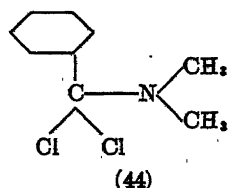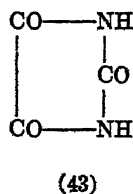
Diazo and diazoamino compounds

Saranew[158] observed that with diazoaminobenzene in benzene solution the passage of phosgene gives a white crystalline compound insoluble in benzene and ligroin. It is slowly soluble in cold water but readily so in hot water which decomposes it with the evolution of nitrogen and the formation of phenol and di-

[155] Basarow. Ber. 5, 477.
[156] Schiff. Ann. 291, 367.
[157] Hollemann. Ber. 9, 846 (1876).
[158] Sarauw. Ber. 14, 2443 (1881).

phenyl urea. This led Saranew to suggest the following for the constitution and decomposition of the compound in question:

$$
\begin{array}{ccc}
\phi & & \\
\text{N}\!\!-\!\!\text{N}\!\!=\!\!\text{N}\cdot\phi & & \text{NH}\cdot\phi \\
\diagup & & \diagup \\
\text{CO} & = & \text{CO} \qquad + \text{N}_2 + \text{C}_6\text{H}_5\cdot\text{OH} \\
\diagdown & & \diagdown \\
\text{N}\!\!-\!\!\text{N}\!\!=\!\!\text{N}\cdot\phi & & \text{NH}\cdot\phi \\
\phi & &
\end{array}
$$
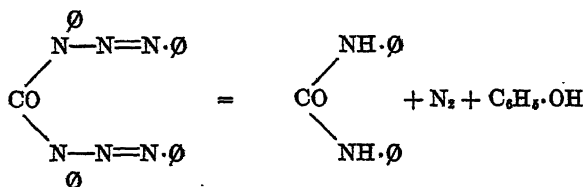
Similar phenomena were observed in the reactions of p-diazo-aminotoluene but diazoaminobenzene-3-carboxylic acid,[159] in the reaction with phosgene gave benzoyl chloride m-oxybenzoic acid and nitrogen. Similar experiments with diazobenzene-p-bromaniline gave no conclusive results.

Amidoximes

By adding powdered phenyloxyethenylamidoxime to a solution of phosgene in benzene Gross[160] obtained a compound in small silvery leaflets m.p. 131°. It is soluble in alcohol and ether but not in water or dilute acids. Gross states that its formula is as in (45), but it is far more likely that condensation takes place through the amino groups, giving the symmetrically disubstituted urea.

$$
\begin{array}{ccc}
\text{NH}_2 \qquad \text{H}_2\text{N} & & \\
\diagup \qquad \diagdown & & \\
\text{CH}\cdot\text{C} \qquad\qquad \text{C}\cdot\text{CH} & & \\
| \quad \diagup \qquad\qquad \diagdown & & \\
\text{OH} \quad \text{N}\cdot\text{O}\!\!-\!\!\text{CO}\!\!-\!\!\text{O}\cdot\text{N} \quad \text{HO} & &
\end{array}
\qquad
\begin{array}{c}
\text{NH}_2 \qquad \text{H}_2\text{N} \\
\diagup \qquad\qquad \diagdown \\
\text{CH} \qquad\qquad\qquad \text{C} \\
\diagdown \qquad\qquad \diagup \\
\text{N}\cdot\text{O}\cdot\text{CO}\cdot\text{O}\cdot\text{N}
\end{array}
$$

$$
(45) \qquad\qquad\qquad\qquad (46)
$$
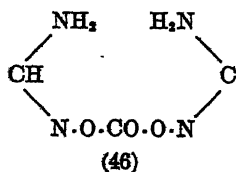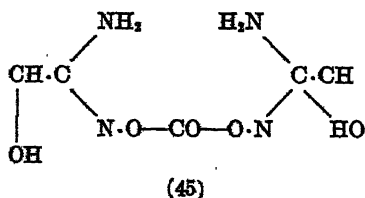
Falk[161] obtained a similar compound from benzenylamidoxime, concerning which a similar uncertainty of constitution exists (46). Exactly the same question of constitution arises concerning the products on the interaction of phosgene and the salt of

[159] Sarauw. Ber. 15, 42.
[160] Gross. Ber. 18, 2480.
[161] Falk. Ber. 18, 2471.

amino-acids containing a hydroxy group. Thus Aloy and Ra-bout[162] observed that phosgene was without action on p-oxy-phenylaminoacetic methyl ester, although on dissolving the latter in alkali phosgene gave a non-crystalline gelatinous condensation product. The authors assumed this to have the constitution (47), but it is probably the symmetrically disubstituted carbamic ester (48), seeing that its preparation requires
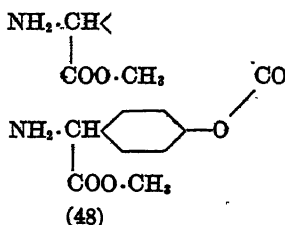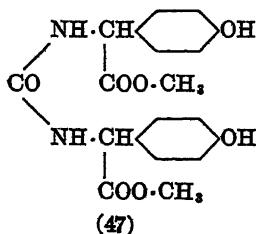
the presence of alkali. A similar compound from tyrosine[163] is also stated to have been obtained in this manner.

Thioureas and thiosemicarbazides

The reaction of phosgene on thioureas and thiosemicarbazides throws some slight light on their constitution, which from the evidence available would seem to comprise a thiol group. Thus, Will[164] found that the reaction between phosgene and diphenyl thiourea proceeded easily and that a compound was formed in small needles m.p. 87. To this compound he arbitrarily assigned the constitution
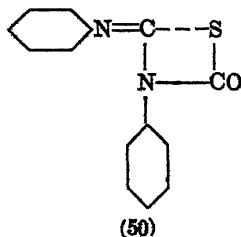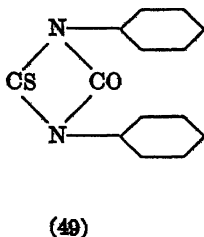
[162] Aloy and Rabaut. Bull. Soc. Chim. (4), 9, 253.
[163] Hugouneng and Morel. C. R. 142, 48 (1906).
[164] Will. Ber. 14, 1486.

(49), but the later work of Wolf[165] has shown that the action of phosgene and thiophosgene on ureas and thioureas falls into line with that of the other chloracid chlorides, and gives rise to a cyclic compound of the type (50). The action of phosgene on thiosemicarbazides has already been discussed.

Thiols

The reaction of thiols or their alkali metal derivatives with phosgene proceeds normally. Thus Saloman[166,167] observed the formation of the compound (51) from sodium mercaptide and phosgene, and was later able to prepare the chlorocarbonyl compound (52) by using phosgene

$$
\begin{array}{ll}
\text{S·Et} & \text{S·Et} \\
\diagup & \diagup \\
\text{CO} & \text{CO} \\
\diagdown & \diagdown \\
\text{S·Et} & \text{Cl} \\
(51) & (52)
\end{array}
$$
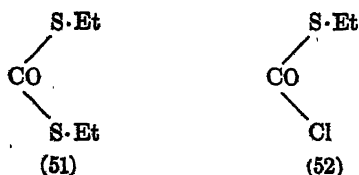
in excess. The corresponding compound from amyl mercaptan was prepared by Schoene[168] as an evil smelling liquid b.p. 190°. It undergoes the usual condensations with ammonia and amines.[169] The reaction is only capable of extension to the xanthates, when thiophosgene replaces phosgene. In this case potassium ethyl xanthate gives the

$$
\begin{array}{cc}
\text{Et·O·CS} & \text{CS·O·Et} \\
| & | \\
\text{S·CS·S} \\
(53)
\end{array}
$$
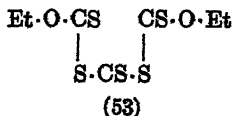
compound (53). If phosgene is used none of the corresponding carbonyl derivatives are formed.

[165] Wolf. Ber. **25**, 1456.
[166] Salomon. J. Pr. Ch. (2), **6**, 433 (1872).
[167] Salomon. J. Pr. Ch. (2), **7**, 254 (1873).
[168] Schoene. J. Pr. Ch. (2), **30**, 416.
[169] Willcox. J. A. C. S. **28**, 1031 (1906).

Phosphines

An isolated reference[170] to the behaviour of phosphines with phosgene shows that when phenylphosphine (the phosphorus analogue of aniline) is reacted with phosgene, three compounds are formed,

$$P \cdot Cl_2 \qquad P=\!\!=\!\!P \qquad \qquad P \cdot CO$$

(55)
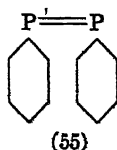
phosphenyl chloride, diphosphenyl, and the phosphorus analogue of phenyl isocyanate (56).

Aldehydes

By reacting aldehyde vapour with phosgene a colourless limpid liquid was obtained by Harnitz and Harnitzki[171] They termed it "chloracetin" and associated with it the formula C_2H_3Cl. Later it was found that the compound reacted with sodium methoxide[172,173] to give acetone—a reaction which, at the time, constituted a new synthesis of that compound,—and which was written:

$$C_2H_3Cl + Na \cdot O \cdot CH_3 = CH_3 \cdot CO \cdot CH_3 + NaCl$$

Later Kekule and Zinke[174] observed that when an aldehyde and phosgene were allowed to react a large amount of paraldehyde and metaldehyde was formed, while it remained for Eckenroth[175] to point out that in addition the chloracetin of the older writers was in reality ethylidene dichloride. Thus:

$$CH_3 \cdot CHO + CO \cdot Cl_2 = CH_3 \cdot CH \cdot Cl_2 + CO_2$$
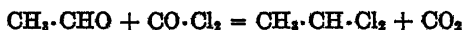
[170] Michaelis and Dittler. Ber. 12, 339.
[171] Harnitz-Harnitzky. Ann. 111, 192 (1859).
[172] Friedel. C. R. 60, 930 (1865).
[173] Friedel. Ann. Chim. Phys. (4), 16, 403 (1869).
[174] Kekule and Zinke. Ann. 162, 125 (1872).
[175] Eckenroth. Ber. 18, 518.

Its interaction with sodium methoxide is then readily explained by the formation of the intermediates:

$$CH_3 \cdot CH \begin{smallmatrix} O \cdot CH_3 \\ \\ O \cdot CH_3 \end{smallmatrix} \longrightarrow CH_3 \cdot CO \cdot CH_3 + CH_3 \cdot OH$$
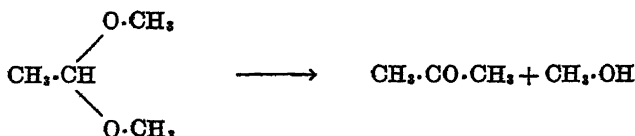
The action of phosgene on aldehydes in the presence of secondary bases,[176] such as quinoline or diethylamine leads to the formation of compounds such as (57) and (58):

$$R \cdot CH \begin{smallmatrix} O \cdot CO \cdot Cl \\ \\ Cl \end{smallmatrix} \qquad R \cdot CH \begin{smallmatrix} O \cdot CO \cdot O \\ \\ Cl \end{smallmatrix} CH \cdot R \begin{smallmatrix} \\ \\ Cl \end{smallmatrix}$$

(57)   (58)
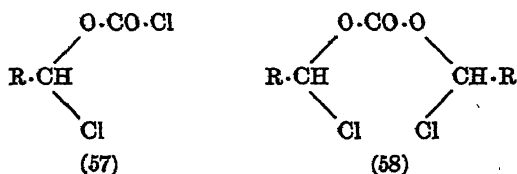
Thus chloral gives a compound $C \cdot Cl_3 \cdot CH(O \cdot CO \cdot Cl) \cdot Cl$ as an oil b.p. 78–80/14 mm.

Nitriles

Henle[177] passed phosgene into acetonitrile and obtained what he thought to be a compound $C_2H_5 \cdot CN \cdot CO \cdot Cl_2$. In all probability the phenomenon was one of mere solution.

Heterocyclic compounds

When pyrrole potassium[178] is suspended in ether and phosgene is passed in, a vigorous reaction takes place and on evaporation of the ethereal solution an oil remains. This oil consists of two compounds di-pyrryl urea which is volatile in steam and can be recrystallised from alcohol in needles, and di-pyrryl ketone which is crystalline but non-volatile.

With pyridine the pentavalent compound (59) is formed

[176] Farbenfab. Bayer. G. P. 121223.
[177] Henle. Ann. 106, 285 (1858).
[178] Ciamician and Magnaghi. Ber. 18, 44.

first,[179,180] although subsequent stages in the reaction have not been fully elucidated.

(59)
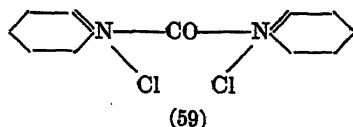
With piperidine derivatives, however, the formation of the urea is more certain. Thus in the case of p-acetylaminobenzylpiperidine observed by Kuehn[181] the action of phosgene gives rise to two compounds p-acetylaminobenzyl chloride and di-piperidyl urea:

$$2 \cdot CH_3CO \cdot NH \langle \;+\; \rangle CH \cdot N(C_5H_{10}) + CO \cdot Cl_2 =$$

$$2 \cdot CH_3CO \cdot NH \langle \;\;\; \rangle CH_2 \cdot Cl + CO \cdot (N \cdot C_5H_{10})_2$$
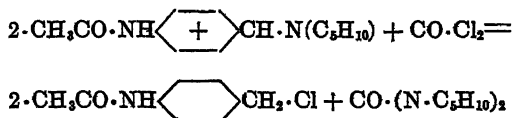
Ketenes

The interaction of diphenylketene and phosgene has been recorded by Staudinger.[182] The reaction is one of simple addition, leading to the formation of an acid chloride of gem-diphenylmalonic acid:
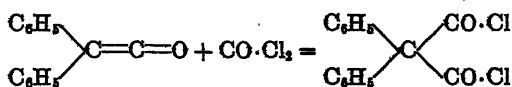
Grignard compounds

Grignard[183] himself investigated the reaction of phosgene on the organo-magnesium halides, and observed that in no case was the ketone formed according to the equation:

$$CO \cdot Cl_2 + 2 \cdot R \cdot MgBr = R \cdot CO \cdot R + 2 \cdot Mg \cdot BrCl$$
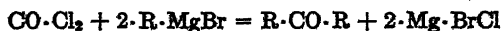
[179] Powell and Dehn. J. A. C. S. **39**, 1717.
[180] Heyden. G. P. 109033.
[181] Kuehn. Ber. **33**, 2900.
[182] Staudinger, Goehring and Schoeller. Ber. **47**, 40.
[183] Grignard. C. R. **136**, 815.

In all cases the secondary and tertiary alcohols were formed according to the equations:

1. $COCl + 3 \cdot R \cdot MgBr = R_2 \cdot CH \cdot O \cdot Mg \cdot Br + R \cdot CH_3 \cdot H + 2 \cdot MgBrCl$
2. $CO \cdot Cl_2 + 3 \cdot R \cdot MgBr = R_3 \cdot C \cdot O \cdot MgBr + 2 \cdot Mg \cdot BrCl$
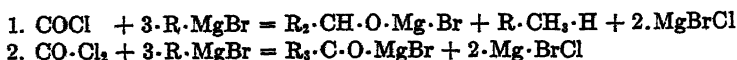
Sachs and Loewy[184] observed that the main product was the tertiary alcohol and that the reaction could be extended to the aryl compounds. Thus, when phosgene is allowed to react on a benzene solution of phenyl magnesium bromide a crystalline product is obtained, which on hydrolysis yields triphenyl carbinol. Tritolyl and tribenzyl carbinols have been prepared in the same manner.

Salts of acetoacetic ester

When phosgene reacts on the sodium salt of acetoacetic ester, the reaction[185] is one of simple chlorination and leads to the formation of monochloracetoacetic ester as a limpid colourless liquid b.p. 192–200°.

$$
\begin{array}{ccc}
CH_3 \cdot CO \cdot CH \cdot COOEt & & CH_3 \cdot CO \cdot CH \cdot COOEt \\
| & = & | \\
Na + CO \cdot Cl_2 & & Cl + NaCl + CO
\end{array}
$$
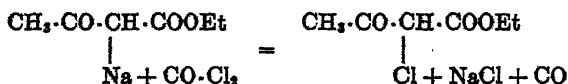
The reaction with the copper salt[186] of acetoacetic ester, however, proceeds to the formation of a dimethyl-γ-pyrone dicarboxylic ester, which is obtained in yellow crystals m.p. 79° to 80°.

The physiological action of phosgene[187,188,189]

The extended use of phosgene in modern warfare as a poison gas lends an excuse for a short note on its physiological action. The chief effect of phosgene is said to be the change in concentration of the blood in animals submitted to its influence. Thus, during the first five to eight hours the blood is considerably di-

---

[184] Sachs and Loewy. Ber. 36, 1588.
[185] Buchka. Ber. 18, 2090.
[186] Conrad and Gutzeit. B. 19, 19.
[187] Hertz. J. Ind. Chem. Eng. 11, 9.
[188] Underhill. Chem. Abs. 1910, 2929.
[189] Samartino. Arch. de. Farmacol. speri. 25, 30.

luted, a phenomenon which is followed by a rise of temperature and at a later stage by the concentration of the blood solids.

In poisoning by phosgene the first symptoms are those of distress and dyspnoea, with coughing and the expectoration of a thin yellow mucous together with occasional vomiting. Later there is severe cyanosis which persists even after the administration of oxygen, death usually taking place from heart failure when the patient attempts some slight physical effort. Post-mortem examination shows that there is much oedema of the lungs, laryngitis, agglomeration of the oesophagus and stomach, and cellular necrosis.

A CHEMICAL CONCEPT OF THE ORIGIN AND DEVELOPMENT OF LIFE[1]

. A PRELIMINARY PRESENTATION

VICTOR C. VAUGHAN

Washington, D. C.

The concept of the origin and development of life which as a result of your kind invitation I am to present to you this evening, has not been evolved from my inner consciousness but has resulted from more than twenty years of experimentation in my laboratory modified by the work of others. In the nineties I was seeking a method by which I might obtain large quantities of some low form of life free from contamination. This quest ended in my devising my large bacterial tanks, with which I was able to secure pure bacterial substance by the kilogram and was able to demonstrate the following fundamental facts.

1. Bacterial substance consists of glyconucleoprotein.

2. It contains no cellulose, consequently bacteria are not plants.

3. Bacterial substance shows no differentiation into cytoplasm, nucleus, or nucleolus and undergoes no mitosis, consequently bacteria are not "cells" as the morphologists would interpret this term.

4. On cleavage with acid or alkali, bacterial substance yields carbohydrates, amino-acids and purine bases.

5. It may be split into poisonous and nonpoisonous portions, with evidence that the cleavage follows definite chemical lines.

6. Dead pathogenic bacterial substance kills animals with the same symptoms and like lesions to those which follow inoculation with the living organism. Therefore the symptoms and lesions

of a disease such as typhoid fever are not due directly to the growth of these bacilli in the patient's body but result from the cleavage of the bacterial substance by some agency supplied by the body of the host.

7. Nonpathogenic bacterial substance furnishes as much poison as does the pathogenic. Therefore, immunity to certain organisms cannot be due to the absence of poison in these organisms but must be explained in some other way.

8. Vegetable and animal proteins such as edestin from hemp seed and casein from milk contain as much poison as do the pathogenic bacilli.

9. All proteins contain a poisonous group. It will be understood that none of these poisons are active when given by mouth and are so only when introduced parenterally. This is because protein cleavage in the alimentary canal is different from that occurring in the blood and tissues.

I may add that the above statements in all essentials have been verified by workers in this country, France and Germany. I published them in book form in 1913. I am not tonight going farther into the details of my experimental work but will devote my time to the conclusions which I have drawn. Even with this limitation I can only present a preliminary outline, awaiting opportunity to write in further detail.

THE ORIGIN OF LIFE

How can we differentiate between non-living and living matter? What is the earliest manifestation of the acquisition of life? Certainly matter does not cease to be matter when it becomes endowed with life. An atom of nitrogen in ammonia is still nitrogen when it is incorporated in a more complicated protein molecule. I can say with much confidence that the conversion of non-living into living matter is accompanied by increased molecular lability. By this I mean that the atoms or electrons within the molecule are energized. Their orbits are enlarged. Within their orbits they move with greater speed. Their chemism is intensified so greatly that they are now able to drag into their orbits atoms and possibly molecules which have hitherto

been beyond their grasp. In other words the molecules begin feeding on outside matter. All living things absorb, assimilate and eliminate. This means that metabolism or trading in energy begins. Such is the first evidence of life. Have we any idea of the nature of these primitive living molecules? Yes. They were and are protein molecules. There is no life save in proteins. These are polymers of amino-acids. The amino-acids, at least the simplest of them, have been and are today being formed under proper environmental conditions from inorganic substances. Furthermore, each protein differs from all others in its content, kind or position in the molecule, of amino-acids. Up to the present time less than twenty of these bodies have been found in nature but with this small number, numberless proteins are formed, much as all the words in our language are formed by varied groupings of the twenty-six letters in our alphabet. Simple proteins yield only alpha-amino-acid on hydrolysis.

In my opinion simple proteins are not living. There must be in the living molecular structure a carbohydrate group thus converting a simple protein into a gluco-protein. I have found two carbohydrates in bacterial substance. One of these which exists in some bacteria to the extent of 10 per cent, I believe to be attached to the nuclein group, while the other is attached to the nitrogen. With this glyco-protein we have a battery and this begins to operate under proper stimuli such as heat, light, electricity or the chemical constituents of something in the medium in which the molecular battery exists. In other words the stimulus is some form of energy. What causes the amino-acids to be synthesized I do not know. Emil Fischer has however synthesized amino-acids and has obtained a product which closely resembles natural protein.

Irritability or re-activity as Ralph Lillie prefers to call it, has long been known as a universal property of living matter. This means that the rate at which energy is received and discharged by the living protein may be altered, increased or decreased, by external stimuli which may be brought into contact with it. The stimulus may be in the form of food or fuel which brings to

the organism, or the living system as the morphologist calls it, energy in the potential form, which is then discharged in the kinetic form. Metabolism is regulated by environment. Reaction between the organism and its environment is essential to all living matter. Without this, life cannot originate or having originated continue indefinitely. One can conceive of a piece of chalk or a lump of carbon existing indefinitely without reacting with its environment, without absorbing or eliminating, but one cannot conceive of a bacterium or a yeast cell retaining life indefinitely under these circumstances. I am fond of repeating a statement first employed I believe by Allen, "Living matter differs from dead in that the former trades in energy while the latter does not."

Still another attribute of living matter is its ability to reproduce itself. As I conceive it, the early forms of life must be particulate, not necessarily as this term is understood by the morphologist, but in a chemical sense, meaning that living matter maintains its molecular identity in no matter what form or environment it exists. The early forms of life must at the same time be small, microscopic or ultramicroscopic because the reaction between the organism and its environment can occur only when the reacting bodies are brought into immediate contact. This holds true whether we consider the lowest or the highest forms of life, whether we subscribe to the cellular or a chemical theory of the origin of life. In man, the highest form of life, contact is just as intimate, the food material being brought to the cells by the blood and lymph. This holds whether the energy be brought to the organism in the potential or kinetic form. There can be no question as to the nature and manner of reproduction in the lowest forms of life since we can see and study it in low forms such as bacteria. Reproduction occurs by fission.

If we assume that there was an Azoic period in the history of the earth, a period in which life even in its simplest forms did not exist, and we must assume this if we will accept the geologist's concept of the origin of the earth, then it follows of necessity that

there was a certain time at which life on earth began. The evidence today indicates that energy derived from the sun is the original source of life. The chief difference between inanimate and animate matter is in its energy content. The forces in the sun's rays have energized dead matter into life. Perhaps, as suggested by Mathews the bulk of this energy is carried by the oxygen atoms within the molecule. As to what was the form of this first life we can but conjecture, but from the evidence, some of which I have just presented you, I believe that we are safe in assuming that life as such did not exist before the evolution of the protein molecule. Every element which makes up the protein molecule exists also in the inorganic chemical world. There was a time when organic chemical compounds did not yet exist. Henry, referring to the possibility of artificial production of organic compounds, wrote, "It is not probable that we shall ever attain the power of imitating nature in these operations. For in the functions of a living plant a directing principle appears to be concerned, peculiar to animated bodies, and superior to and differing from the cause which has been termed chemical affinity." And yet only a short time after this Wöhler succeeded in synthesizing urea.

Moore and Webster appear to have succeeded in synthesizing formaldehyde from carbon dioxide and moisture under the influence of ultraviolet rays and in the presence of an inorganic colloid. It matters not whether as has been suggested by von Baeyer, formaldehyde was the first organic substance produced in nature leading toward the development of life. The point is that it has been shown that organic substances may be synthesized in the laboratory from inorganic substances and that such simple organic structures as amino-acids may be synthesized experimentally into compounds closely resembling natural proteins. It makes no difference whether we can now or ever will be able to reproduce each and every step to the ultimate development of life. Failure in no way invalidates our hypothesis any more than our inability to build a star or planet disproves existing views as to the probable structure of the universe.

At some stage in the evolution of life the cell as we know it

today came into existence. It is back to this point that the morphologist traces the origin of living matter and beyond this he does not allow himself to go. The doctrine "omnis cellula ex cellula" may perhaps hold after the first cell came into being but the chemist cannot accept the cell as the lowest or the original manifestation of life. Nearly twenty years ago I first stated my belief that life is fundamentally chemical and may, indeed probably does exist in simpler and less tangible forms than the living cell or even the living bacterium which I do not regard as a true cell for it contains no differentiated cytoplasm and nucleus.

It becomes incumbent upon the chemist who denies the contentions of the morphologist to explain how the cell may evolve from simpler forms of life. This we cannot yet do but the work of DuNouy on the surface equilibria of colloids opens interesting fields for speculation. This author presents evidence that the most probable configuration of equilibrium in a protein colloid solution is in the form of a cell. This would present a minimum of free energy compatible with the total energy. If a microscopic droplet of protein solution is sprayed into the air the constituent molecules of this droplet will proceed to arrange themselves in thermodynamic equilibrium, with relation to each other. In this process the droplet will become coated with a surface layer of protein three hundred or four hundred times more viscous than the interior. If the droplet has a diameter of ten microns, equilibrium may be established within four seconds. If the diameter is but three microns, equilibrium is established in about one second. If equilibrium has been established before the droplet completes its fall the concentrated surface layer will be strong enough to maintain its shape even if it strikes a dry surface. The presence of carbon dioxide or hydrochloric acid gas or ultraviolet rays suffices to render some of the constituents of the protein layer insoluble, thus enabling the droplet to keep its individuality even though it fall into pure water. Assuming as I have stated before that energy and its transformation is one of the dominant characteristics of life, we have evidence in the work of DuNouy that the cell may be but the logical consequence of the tendency of the protein molecule or molecules to establish dynamic equilibrium.

What is the smallest form of living substance known?   The smallest living structure known today is that entity which has been described and studied in greatest detail by d'Herelle and to which he has given the name of bacteriophage.   This living particulate chemical substance is much smaller than the smallest known cell and bears out my hypothesis first stated nearly twenty years ago.   d'Herelle gives to the bacteriophage the generic name protobe or first life.   It is without doubt the simplest form of life known today but I regard it as not proven that the first life was not even simpler.

The bacteriophage fulfills all the criteria of life.   It can assimilate in a heterologous medium, transforming a heterologous substance into homologous bacteriophage substance, a substance distinctively its own.   With this function of assimilation it also possesses the function of adaptation to changing environment. Furthermore it possesses the faculties of reproduction and variability.   The substance is antigenic, has the chemical constitution of protein, possesses as great and prolonged viability as bacterial spores and appears to be an electro negative colloid just as are the majority of the bacterial species.   The dimensions of the bacteriophage corpuscle are approximately those of the serum globulin micella, its diameter being about twenty millimicrons.   The substance appears to be thermolabile, its virulence being destroyed at about 75°.   The protein micella is the colloidal unit.   It is the smallest possible particle of matter in the colloidal state.   Possibly as d'Herelle states it is the unit of living matter and cells are constituted of a union of micellae. The bacteriophage is of about the size of a micella.

<center>BACTERIA</center>

I do not regard bacteria as the simplest form of life.   Their chemical structure is very complicated.   They are essentially nucleins and their chief function is to multiply.   Whether the individual consists of a single or many molecules I do not know. Probably their structure is multimolecular but if so the chemism between the molecules must be very strong.   I know of no way of distinguishing between intermolecular and intramolecular activity.

Bacteria will live under most diverse conditions. They will grow in a medium which contains organic nitrogen only in the form of ammonia. They continue to live or at least to retain life under wide ranges of temperature. When food is scarce they go into a resting or spore stage. They multiply by fission. In them acquired characters such as increased or decreased virulence are transmitted. They are antigenic and can be shown susceptible to classification in groups by their antigenic reactions.

While the bacterial cell is morphologically simple in structure, it is as complex in chemical composition as are the cells of the animal body. I know of no work done since I reached this conclusion which throws any doubt upon it. The conclusion that I would draw therefrom is either that bacteria are already relatively high up in the scale of life or else that even the simplest forms of life consist of relatively complex aggregations of protein molecules.

The general constancy and immutability of bacterial types is illustrated in the history of epidemic disease. Generally speaking these diseases run true to type through their recorded history, be this short or long. Tubercle bacilli found in Egyptian mummies present the same characteristics and cause the same type of tissue destruction as do tubercle bacilli in the consumptive of today. The characteristic symptoms and lesions of smallpox observed and described by Indian writers before the Christian era show no essential variation from those which manifest themselves in the unprotected individual of today. Through all the centuries there has been no important mutation in the smallpox virus, nor any marked modification in its behavior when introduced into the human body. The most ancient descriptions of the plague are so plainly indicative of the disease as we know it in the present generation that there can be no mistake of the identity of the virus of this disease in most ancient times with that of the present. The pneumonias of today are marked by the same seasonal variation, characterized by the same modes of onset, by like avenues of progress, and by similar results with those seen and described by Hippocrates. Because bacteria and protozoa are low forms of life it has been assumed that they are

especially liable to marked mutations involving alterations in chemical composition, and what is of more importance, so far as pathogenic organisms are concerned, in their effect upon man. In my opinion the assumption that bacteria and protozoa readily undergo mutation is not warranted by any facts which can be gathered in a study of the history of infectious diseases. I am ready to assert that there has been less mutation in the tubercle bacillus or the virus of smallpox since the beginning of recorded time than there has been in man and the other higher animals.

We do not know the nature of the filterable viruses such as that of smallpox but it is possible that they are of the nature of protobes such as d'Herelle's bacteriophage. There is some evidence that as time goes on we will be able to establish a more definite connection or association between the protobes and bacteria. In the case of the tubercle bacillus for instance there is evidence that ultrafiltrates of the tubercle bacillus contain what appears to be a living virus and the evidence suggests that this is a small granular form of the tubercle bacillus. As to whether this is a tubercle bacillus micella we can only surmise.

THE CHEMISTRY OF PARASITISM

Having outlined my concept, a chemical concept, of the early development of life I desire now to present to you my interpretation of the manner in which life became differentiated into its many forms. My understanding of this complex phenomenon, as I have said, is not based upon pure philosophic induction but upon experimental observation in my own laboratory. Before discussing the origin of species therefore, I must summarize briefly my conception of the life processes as I have observed them in bacteria.

The chief function of life is self perpetuation. If this function is to continue active, the living substance must be so situated that it can procure nutritive material from its immediate environment. While energy is furnished in the available carbohydrates and fats and while water and certain minerals are requisite, the structural and reproductive requirements of the protein molecule are met only by protein material, of which the

basis is the amino-acid. All living substances are proteins. The nature of the protein differs for every different type of life. This difference is due to variation in the number, nature or intramolecular arrangement of the constituent amino-acids. There may be other differences which present methods have not as yet enabled us to recognize. If the necessary pabulum were always available as pure amino-acids and in the correct proportions for the particular living cell, the matter would be simple. However, as a rule the available organic food supply consists of combinations of amino-acids in varying degrees of complexity, up to the complete protein molecule.

In order that the living substance, let us say a bacterium, may assimilate this food it first becomes necessary to disrupt the heterologous protein molecule into its constituent amino-acids so that these may be absorbed and built up into the bacterial structure. Bacteria secrete enzymes or ferments for this purpose. So do all living cells. These ferments will digest certain proteins but not all proteins. If a living cell is in contact with a foreign protein against which it does not possess a digestive ferment it will gradually evolve a ferment specific for that protein. I believe it to be a fundamental law that a living cell in contact with a foreign protein will evolve an enzyme to destroy that protein. Many years ago Duclaux showed that penicilliun glaucum grown on starch produces invertase only. On lactose it produces lactase in addition. On milk it elaborates a proteoclastic enzyme. The ability of living cells to produce specific enzymes to meet the necessity for disrupting the substrate with which they come in contact is essential to existence both under normal and abnormal conditions. It enables the cells to feed upon assimilable substances and to destroy injurious ones. The ease with which living cells may function in these directions is dependent upon many and varied factors such as temperature, physical and chemical conditions, the activity of the cell which is seeking to feed or protect itself and the constitution of the body upon which it acts. Here lie many problems awaiting future investigation.

Much remains to be learned of the nature of ferments or

enzymes. They are particles of matter, some of them wholly simple like spongy platinum, others highly complex like the yeast ferment or pepsin or trypsin. Enzymes are inanimate storehouses of energy which may be brought into action under proper environment or on coming under the influence of certain physical or chemical stimuli. They may be compared roughly to storage batteries. Ferments may be protein but are not animate.

In the same way that the bacterium will elaborate a digestive enzyme, so also will the body cell do this when it comes in contact with a foreign protein. The typhoid bacillus on entrance into the body grows luxuriantly during the incubation period, for in the blood it finds an abundance of available food material in the same simple form that is available for the body cells. Living typhoid bacilli have been found in the blood of man during the incubation period, before any symptoms of the disease have become manifest. The germ is converting body proteins or at least the amino-acids of the blood into typhoid bacillus protein. The reaction is synthetic and there are no symptoms. But the body cells have been stimulated by the presence of a foreign protein and in about ten days they have elaborated a ferment or enzyme which will break down this foreign protein. As soon as this defense reaction becomes well developed disease becomes manifest. Now, typhoid bacilli are being destroyed, the process is analytic, the protein poison is being liberated.

During the incubation period the process is constructive. After the body cells have learned to elaborate a ferment which will destroy the typhoid bacillus the process in turn becomes destructive and in this destruction the protein poison appears to be liberated. This poison I have found to be present in every protein which I have so far examined. It exists not alone in pathogenic bacteria but also in the nonpathogenic and even in such otherwise innocuous proteins as egg white and the proteins of the cereal grains. Indeed edestin from hemp seed and casein from milk, and egg white furnished me the largest and most satisfactory amounts of protein poison. While I have been unable to obtain this substance in anything approaching a pure state, it appears to contain many amino-acids and, apparently

should be classed as a polypeptid. It is only poisonous when administered parenterally, for alimentary digestion apparently further breaks it up into the simple amino-acids. In the case of typhoid fever which I have used as an example the symptoms result from the liberation of the protein poison. The severity of the disease depends upon the amount and rapidity of liberation of the poison. The very small doses of the poison which will produce serious symptoms experimentally indicate its possession of a high degree of energy. In my opinion it kills by tearing off from certain body cells secondary and functioning chemical groups.

The perpetuation of life depends upon the ability of living substance to convert heterologous proteins into homologous proteins. This holds equally for the higher forms of life for if in the case under consideration the human body is unable to convert typhoid proteins into human protein, the result will be disastrous. True, this conversion is of itself not without danger.

The ability of a bacterium to produce disease after it has entered an animal depends mainly upon two factors. First it must be able to establish for itself a parasitic existence in its host. It must be able to sustain itself and multiply its kind on the pabulum within its reach. Second, there must be no destructive enzyme already existing in the body tissues, against this particular bacterium. Disease depends in great part upon how abundantly a given microörganism may multiply in the tissues before the body cells have completed the elaboration of a destructive enzyme. Where one is already in existence the bacterium is destroyed at once and only an infinitesimal amount of poison is liberated. No symptoms result. The seriousness of the symptoms depends upon the amount of poison liberated and the rapidity of its liberation. Of course, there are other factors in certain diseases such as location within the body, the secretion of a toxin by the bacterium and the like.

THE ORIGIN OF SPECIES

Let us now take up a consideration of those factors which may have had a bearing in the origin of species. At this point

I am not so interested in the inheritance of identical character-istics as I am in the inheritance of altered or acquired charac-teristics for it is by virtue of the latter that new species develop. I find no great difficulty in understanding that living substance might readily reproduce itself in its entirety but I am highly interested in the intimation that it can produce another living substance different from any that has been known previously.

I have said that the characteristics of bacteria have remained remarkably constant throughout the history of disease.   There is an exception, one which I believe to be of fundamental im-portance in the development of species.   I have said that the tubercle bacillus as it occurs in man appears to have undergone no remarkable change through recorded history.   This is quite true but at the same time there is a tubercle bacillus which infects fowls which is not quite the same germ and yet another which infects cattle.   Dr. Calmette has at the Pasteur Institute in Paris a strain of tubercle bacillus which he has cultivated artificially for over thirteen years and which appears to have lost entirely its ability to infect man and the lower animals.

The constancy of bacterial types and indeed of all living sub-stances depends upon a relatively unchanging environment.   In the lower forms of life environment has a very definite influence upon the characteristics of life.   Furthermore, alterations in the structure of the protein molecule resultant on environmental changes may be and are inherited.   A microörganism living in a milieu in which the pabulum is readily assimilated and trans-formed into homologous proteins will thrive.   If, on the other hand, the environment is one in which the available food material is of widely different constitution from that of the living sub-stance, continued existence will depend upon the ability of the microörganism to elaborate a ferment capable of disintegrating the foreign protein or protein-like substance into its constituent amino-acids so that they may be available for assimilation.   If some of these amino-acids are deficient in quantity for the par-ticular living substance, continued existence will now depend upon the ability of the living structure to adapt itself to this deficiency.   If such an adaptation is made, there will result a change in the make-up of the living protein molecule.

While I am emphasizing chemical factors I am not unmindful that physical and other factors also play a part. I can readily understand why many species of animals and of plants have disappeared. No species can continue when it ceases to receive and utilize energy from its environment. A change of a few degrees in the annual average temperature might change markedly the flora and fauna of the area in which it occurs. Climatic factors are more readily recognized in the higher forms of life, but I shall continue to limit myself to the effect of changes in the chemical environment on the lower forms of life.

The presence and availability of new or different amino-acids or similar protein radicles will ultimately determine an alteration in the constitution of the living protein molecule. If this alteration in environment is permanent the altered constitution of the living molecule will likewise become permanent and will remain so as long as the environment is relatively the same. The development of new species in the lowest forms of life depends upon physico chemical alterations in the environment. The persistence of new species so formed is dependent upon the permanency of the environmental changes.

A streptococcus highly pathogenic for the horse will on repeated passage through a laboratory animal such as the mouse or rabbit or guinea pig, gradually lose its high virulence for the horse while acquiring an increased invasive power against that animal through which it is being passed. That there is an actual change in the chemical constitution of this streptococcus is indicated by the fact that after several passages its antigenic power as a horse streptococcus which was originally of high titer becomes completely lost.

Not only this but it has been found that cultivation of a streptococcus in an artificial medium containing the blood of some laboratory animal increases the virulence of the streptococcus against this particular animal. Furthermore the antigenic characteristics of this streptococcus are altered after repeated growth on these special laboratory media. Thus we may speak of a horse streptococcus, a mouse or rabbit or guinea pig streptococcus all derived from the same ancestor, each of them still a

streptococcus, but definitely altered in chemical nature by their immediate nutritive environment.

d'Herelle believes that there is but one bacteriophage but that this like the streptococcus just described is capable of adaptation to growth in a wide variety of bacterial hosts. The Shiga bacteriophage and the Staphylo-phage differ from each other in their predilection, one for the dysentery bacillus and the other for the staphylococcus. These are their foods of choice and they find it difficult to grow on other bacteria. However, adaptation may be accomplished and it is possible to change the Shiga-phage into the Staphylo-phage and vice versa. There is an almost limitless possible number of bacteriophages dependent upon the degree of invasiveness for different bacteria but the evidence presented by d'Herelle would indicate that this is a matter of adaptation to the environment on the part of a single original bacteriophage.

This adaptation is so complete that it involves an alteration in the chemical structure of the bacteriophage which can be recognized in changes in its antigenic properties. Shiga-phage actually becomes chemically different from the Staphylococcus bacteriophage. Alterations in the environment have produced a new species which will maintain its identity as long as the environment remains essentially unchanged. Further environmental changes will produce yet other alterations in the structure of the living molecule, not necessarily a reversion to the original structure but perhaps with the development of an entirely new and more complex structure to suit the requirements of the altered nutritive environment.

Many of the higher forms of life contain two or more proteins no one of which can be said to be more specific than the other for that particular living substance. Such a simple plant as wheat for example contains gliadin, globulin, glutenin, proteose and leucosin, five proteins in all. Now wheat glutenin appears to be similar in its chemical constitution with the glutenin found in barley and in rye. I would interpret this as highly suggestive evidence that wheat, barley and rye evolved from the same primordial ancestor. Environmental changes, possibly variations in

the nutritonal resources, have been responsible for the differentiation of these three grains. The farmer of today knows well the importance of environment. With the same seed, the same heredity, he does not anticipate an equally good or abundant crop in every field in different years. The fertility of the soil, the amount of sun and rain and many other environmental factors play a most important part. If some of these factors are disadvantageous to the continued existence of a grain, this grain must either adapt itself to changed conditions which it will do with alterations in its own structure and appearance, or it must · eventually die out.

In forms of life such as those which we have just been discussing, in which two or more different proteins exist together, we must conceive as possible that species differentiation does not necessarily entail complete change in any or all of the constituent proteins but that in these higher forms new proteins may be added, possibly by differentiation of the original with the result that the new protein and the original both exist in the same living substance. Where a protein has at last been evolved which best fits the functional needs and where its environment remains little changed, its chemical constitution will remain remarkably constant. It is said that the proteins of the lens of the eye are different from the other proteins of the body but are identical in the lenses of a wide variety of animals. Here there is little environmental change for the environment is not the outside world but the blood and lymph.

I have mentioned changes in antigenic properties as indicating alterations in the makeup of the protein molecule. The question might be raised as to whether such changes do necessarily indicate alterations in the constitution of the molecule. The term antigen is employed by immunologists to designate those substances which when introduced into the animal body parenterally lead to the elaboration within the treated animal of a substance which antagonizes or tends to neutralize its own action. Up to the present the weight of evidence is all that antigens are proteins. Moreover it appears that each protein leads to the elaboration of a specific antibody. Thus an animal treated with

the venom of a certain species of snake produces an antibody to this venom and not to the venoms of other species of snakes. The toxin of the diphtheria bacillus produces a diphtheria antitoxin and this has no antagonistic action on other toxins. Each antigen acts specifically and the nature of the antibody formed is strictly specific. The body cells appear to elaborate these antibodies and they do it for self protection. Some of the antibodies neutralize their specific antigen by combining with them and thus rendering them inert. This seems to be true of the antitoxins of diphtheria and tetanus also for vegetable toxins such as those of abrin and ricin and the venoms of snakes. In other cases the antibody renders its antigen inert by disrupting it into its harmless constituents. Could there be better or more conclusive evidence of the ability of the body cells to adapt themselves to their environment and to protect themselves against threatened destruction? Living cells are capable of being trained or educated. In other words their behavior may be modified by changed environment.

It has been shown that the specificity of antigens is dependent upon their chemical composition. For instance there are in milk four chemically distinct proteins and each is capable of causing the body cells to elaborate its own specific antibody. In egg proteins there are three chemically distinct proteins some of which are common to the eggs of different species of birds while others are found in a single or in a limited number of species. It seems to be true that the specificity of an antigen is determined by the location of some aromatic radicles within the structure of the protein molecule. When proteins are hydrolyzed they lose their antigenic properties. My students and I showed many years ago that gelatin which is a hydrolyzed protein and devoid of certain aromatic radicles such as tyrosin and tryptophan, and which contains only a trace of phenylalanin is not an antigen. Like results have been obtained by subjecting true proteins to the cleavage action of digestive ferments such as trypsin and pepsin. Likewise the protamins, which are complexes of diamino-acids, and wanting in the amino-acids, are not antigenic. Free amino-acids are not antigenic. All anti-

gens are colloids, all apparently are proteins. Now, it has been found that certain chemicals, as formaldehyde, nitrous acid, and iodin, may be introduced into the protein molecule without destruction of its antigenic properties, but the antibodies elaborated are specific to the altered proteins and not to the original substances. Some years ago Obermeyer and Pick found that the serum of rabbits treated with proteins which had been radically changed by being iodized or nitrified did not precipitate the native protein but did act upon the altered protein with which the animal had been treated.

All life is protein and the development of new species is due to molecular re-arrangement in the structure of the protein molecule. Something is added or subtracted, or chemical groups within the molecule are rearranged. The recently discovered facts demonstrated by the precipitin and sensitization tests make this certain. By these, proteins may be positively identified either when mixed or unmixed with other proteins. Group relationship may be shown by these methods and up to the present time in no other way. Especially is this true when the results of these tests are measured quantitatively. The proteins of the hen's egg sensitize guinea pigs to themselves and to a lesser degree to the proteins of the eggs of other birds. The proteins of man's blood sensitize animals to themselves and less perfectly to those of the blood of the anthropoid apes. Wheat, rye and barley all come from a common plant and under different environments have developed into three species. In this way varieties and species come into existance.

EVOLUTION IN THE HIGHER FORMS OF LIFE

From the lowest to the highest forms of life environment plays a part of greater or less significance in the development of species. These environmental factors may be chemical or they may be physical. I have presented to you my concept based upon the simpler forms of life for here that very simplicity facilitates more accurate study and interpretation.

In calling to your attention the primary importance of environment in the development of life and the differentiation of species,

it is in no way my desire to intimate that I am not in accord with the prevailing doctrines of heredity. The discussion is along quite different lines for in the latter our interest is in phenomena in which gross alterations are conspicuously absent while in the former it is the alterations which are of chief importance and interest. The genes about which students of heredity are saying much I can accept, if I am permitted to regard these genes as atomic groups, some right handed, some left handed, in the specific protein which reproduces itself.

But I find no difficulty in recognizing the action of chemical environment even in the highest forms of life. Morphologists stress the stability of germ plasm but some of them do admit that certain poisons such as alcohol, lead, mercury, and syphilis may deleteriously affect the reproductive cells. In my opinion, even more striking examples might be given. A boy and a girl born of healthy parents and raised to maturity under normal conditions may migrate into a goiterous district and after acquiring goiters may marry. Their children may be cretins. In this case it is the *absence*, according to the now accepted belief, of iodin in the food and drink which leads to this deterioration. Please understand that it is only the absence of one chemical element which causes this disaster. I am in favor of eugenics but I cannot forget that environment as well as heredity must be taken into consideration. The claim that the reproductive cells are not influenced by the somatic cells is one which I believe to be unwarranted. In seeding it is well to select sound grain, but the harvest will not be determined wholly by this, but will depend to some extent on the fertility of the soil.

What is the optimum relationship as between the chemical environment, particularly the food supply, and the living structure? Without considering other factors that undoubtedly play a part I would say that the more closely the heterologous protein resembles in its make-up the homologous living protein, the more nearly identical its content and proportion of the different amino-acids and associated radicles, the more constant will be the composition of the living molecule and the higher will be the degree of perfection which it will attain while remaining

essentially unaltered chemically. I recognize that there are other essential requirements such as vitamines and the like but the basis of life is protein and in this thesis I have limited myself almost entirely to the consideration of the chemistry of the living protein molecule.

Some living forms such as bacteria feed upon other living forms. They can do so because they can convert and assimilate without difficulty the protein molecule of the host. I consider it possible that the more nearly the proteins of the host resemble chemically those of the invader the greater will be the pathogenicity of the latter. I suggest that some investigator study by the antigenic reactions, the protein relationships between parasite and host. In the case of bacteria, feeding upon man or animal, the objection might be raised that they derive their sustenance, not from the living molecules of the animal, but from the simpler protein food radicles and cleavage products present in the circulating blood and lymph. But this objection cannot hold in the case of test tube experiments in which the available pabulum consists of the tissues of the host.

One might infer that I believe that a cannibalistic existence would be the ideal form of life. But curiously enough even in such low forms of life as bacteria and bacteriophages, cannibalism appears not to exist. Protein molecules endowed with attributes of life, while apparently bent upon the destruction of other forms of life, particularly simpler forms, appear incapable of destroying living substances of identical or nearly identical chemical constitution. This is readily understandable. Where there are two such living substances in apposition, their chemisms would be identical, their spheres of influence the same, their tropisms would balance one another and the result would be no chemical reaction. A solution of ammonium chloride mixed with another solution of ammonium chloride remains ammonium chloride.

Late in the 18th century Lavoissier, scientist, patriot, martyr, showed the process of respiration in man is comparable to the burning of a candle. About 100 years ago, Wöhler made urea synthetically. A few years later Dunglisson and Emmett, in their scantily supplied laboratory at the University of Virginia

announced that the free acid in the gastric juice of man is hydrochloric acid. Dumas in France, Liebig in Germany, and others continued to develop physiological chemistry. About the middle of the last century leading universities in this country provided chairs in this subject. For many years Chittenden at Yale was the standard bearer and on his retirement his good work was continued and amplified by Mendel. Splendid work in this subject has been done by Van Slyke, Lusk, Folin and others. Ehrlich and Hatta, after more than 600 attempts built up arsphenamine synthetically and this with its cogeners has done much to mitigate the plagues of syphilis and allied diseases.

Starling and others have discovered hormones and the brilliant results obtained by Banting in his discovery of insulin are well known. Abel not only discovered epinephrin, but determined its structural formula and it is now made synthetically. The same talented investigator appears to be on the high road to similar results in the study of insulin.

However, all these are inanimate substances and up to the present time no chemist has awakened dead matter into life. It may be that this will never be done. Whatever may be individual opinion on this subject, past, present and even future, failures should not prevent us from interrogating nature and learning so far as possible how she in her great laboratory with boundless facilities and with countless ages in which to operate has accomplished this great result. Without predictions as to what degree of knowledge future researches will reveal I have ventured to present my views on this subject. Should they, even in part, be confirmed the morphologist must radically change his teachings as to the relative importance of heredity and environment. I hold that the lowest forms of life have come into existence through chemical agencies and that environment has been a stronger factor in the evolution of life and in the development of the varieties and species than is believed by the biologist of today.

REFERENCES

VAUGHAN, VICTOR C.: Protein Split Products in Relation to Immunity and Disease. Lea & Febiger, 1913.

D'HERELLE, F.: The Bacteriophage and Its Behavior. Williams and Wilkins, 1926.

DUNOUY, P. LECOMTE: Surface Equilibria of Biological and Organic Colloids. Chemical Catalog Co., Inc., 1926.

LEREBOULLET, P., AND LELONG, M.: Tuberculosis, Paris Medical, 1926, vol. 16, p. 1, January 7.

LILLY, RALPH S.: Protoplasmic Action and Nervous Action. University of Chicago Press, 1923.

MATHEWS, ALBERT P.: Some General Aspects of the Chemistry of Cells. General Cytology, University of Chicago Press, 1924.

THE STRUCTURE OF THE METHANE MOLECULE[1]

VICTOR HENRI

Laboratory of Physical Chemistry, The University of Zurich, Switzerland

During the last ten years a number of different methods of investigating the structure of molecules have been worked out and an almost completely new science of molecular physics has been developed. The results which have been obtained are of very general importance and the application of the theories and principles which have been derived from these studies marks a new period in the development of chemistry.

The structure of the molecule of methane is of the utmost importance to all organic chemistry involving, as it does, fundamental conceptions concerning the carbon atom and an application of these modern methods of investigation to the study of the methane molecule has given results of far-reaching significance.

Methane is a colorless, odorless gas; it shows slight deviation from the simple gas laws and can be liquefied at $-164°C$. It is the most important constituent of natural gas, some samples of natural gas showing as much as 99.3 per cent of methane. Since the work of Pasteur, van't Hoff and LeBel more than fifty years ago there has been very general acceptance of the *tetrahedral structure* of the methane molecule, the carbon atom being at the center of the tetrahedron and the four hydrogen atoms at the corners (fig. 1).

From the standpoint of the present-day ideas of atomic structure two of the six planetary electrons of carbon are near the nucleus and the other four are in outer orbits. The hydrogen atom has one planetary electron and the electronic structure of the methane molecule may be shown schematically by figure 2 where the dots represent electrons, the inner electrons being inclosed within the circle.

The evidence upon which the tetrahedral structure of the methane molecule is based depends primarily upon the fact that methane does not form isomeric derivatives. For example, there is only one dichlor-methane whereas, if the structure were not a tetrahedron there should be two isomeric dichlor-methanes. The three general assumptions upon which the whole of organic chemistry has been built are:

1. The equivalence of the four valencies of carbon
2. The tetrahedral structure of methane
3. The rigidity of the structure of molecules

With the development of the science of molecular physics during the last ten years physicists have introduced new methods of
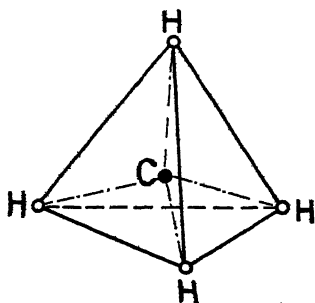
FIG. 1
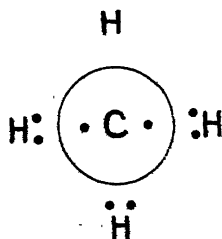
FIG. 2

investigating the structure of molecules and have made important contributions to the theoretical interpretations. These new methods were first controlled by the study of different simple molecules, such as, nitrogen, hydrogen, chlorine, hydrogen chloride, carbon monoxide and then more complicated molecules were examined and now more than seventy different molecules have been studied, of which about fifty have been investigated in my laboratory.

The results of the study of the methane molecule may be briefly summarized as follows:

1. The molecule of methane is *not* a tetrahedron but has the structure of a pyramid. The distance between two hydrogen atoms is equal to 1 Å or 10^{-8} cm. and the height of this pyramid is

equal to 0.37 Å. The relative positions of the hydrogen and carbon atoms are shown in figure 3.

2. When the four hydrogen atoms of methane are replaced by four chlorine atoms the structure changes and becomes a tetrahedron so that the molecule of carbon tetrachloride is tetrahedral. The dichlor-derivatives of methane have a tetrahedral structure

$$a = 0.37 \text{Å} \quad b = 1 \text{Å}$$
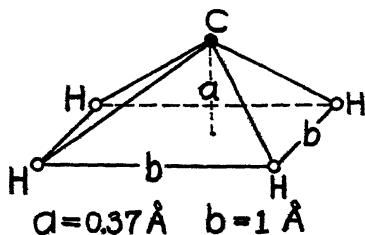
FIG. 3

whereas certain of the tetra-derivatives of methane have a pyramidal structure. For example, pentaerythritol, where each hydrogen atom is substituted by a (CH_2OH) group has a pyramidal structure as has also the molecule of tetra-phenyl-methane.

3. The four valencies of the carbon atom are *not* equal. Two
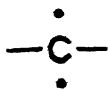
FIG. 4
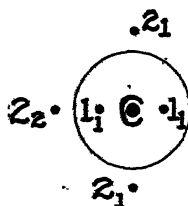
FIG. 5

of them are of one kind and two of another (fig. 4). Our experimental evidence leads us to believe that the six electrons of the carbon atom are distributed in three shells. The two inner electrons are in the 1_1 orbit and of the four external electrons two are in the 2_1 orbit and two in the 2_2 orbit. This relationship is shown in figure 5.

These three conclusions are in absolute contradiction to the classical theories upon which organic chemistry has been built and in the light of this new experimental evidence there should be a complete revision of the chemistry of carbon compounds. It is with pleasure that I acknowledge the very important contributions made by American physicists to the science of molecular physics.

The experimental methods by which these important results have been achieved may be classed into five different groups:

1. The scattering of light
2. The absorption spectra
3. The structure of crystals as determined by x-rays
4. The calculation of the potential energy of molecules
5. The emission spectra of atoms and molecules

These various methods will be examined briefly and the results summarized.

1. THE SCATTERING OF LIGHT

The very familiar example of the effect of the scattering of light is the production of the blue color of the sky by the scattering of the sun's rays by the molecules of the atmosphere. Tyndall (1) was the first to investigate thoroughly the scattering of light and the phenomenon has since been called the Tyndall effect. J. W. Strutt (Lord Rayleigh) (2) developed the theory by which it is shown that when a beam of light of intensity I_0 falls on a gas inclosed in a tube, a portion of this beam is scattered by the molecules of the gas and the intensity of this scattered light, I_1, (fig. 6) in one direction is dependent upon the kind of molecule, the volume of the molecules and the number of them per cubic centimeter. This amount of scattered light is very small, varying from two to five millionths of the intensity of the incident light. If the molecules are isotropic the scattered light is totally polarized, the plane of polarization being determined by the directions of the two beams I_0 and I_1.

During the last seven years a number of experimental investigations of the scattering of light by various sorts of gases have been made by R. J. Strutt (Lord Rayleigh) (3), Cabannes (4),

Gans (5), and Raman (6) and these results have shown that for
the rare gases, helium, neon and argon which are monatomic
there is a total polarization of the scattered light. These molecules
are all isotropic. In the case, however, of the diatomic and
triatomic gases, nitrogen, hydrogen, oxygen, carbon monoxide,
hydrogen chloride, carbon dioxide, water etc., there is only a
partial polarization produced in the scattered light. This in-
complete polarization is due to the anisotropic character of these
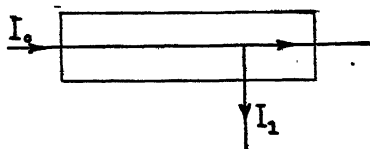
FIG. 6
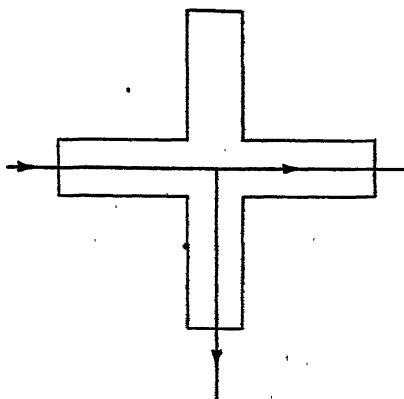
FIG. 7

molecules. A complete theory has been derived by Gans (5)
according to which it is possible to calculate the degree of aniso-
tropy of a molecule from the proportion of the scattered light
which is non-polarized. It is thus possible to show that the
structure of these different di- and triatomic molecules can be
represented by an ellipsoid. A special study of the molecule of
methane and of other hydrocarbons was made by Cabannes (7).

The actual measurement of the amount of scattered light

which is polarized involved great difficulties in technique. The
pure gas was dried and filtered so that it contained no dust and
was then introduced into a tube having the form of a cross (fig. 7).
An intense beam of monochromatic light was allowed to fall on
the gas and a photograph made at right angles to the plane of the
incident light. The amount of polarized light was determined
by a system of nicols. The results obtained with methane showed
that there was a considerable proportion of non-polarized light
thus showing that the methane molecule is anisotropic. This
result excludes the possibility of a tetrahedral structure inasmuch
as the molecule must have two different moments of inertia. In
the case of carbon tetrachloride Cabannes found *isotropic*
molecules.

2. THE ABSORPTION SPECTRA

When a beam of light containing all wave-lengths from the
infra-red to the ultra-violet is passed through a gas or vapor

$$I_a \qquad\qquad\qquad I_1$$

FIG. 8

contained in a tube with quartz ends, analysis of the beam of light
I_1 (fig. 8) after passage through the gas shows that certain wave-
lengths are missing or have been absorbed by the gas. If the
beam I_1 falls on the slit of an infra-red spectrometer or of a
spectrograph the absorption spectrum of the gas or vapor may be
determined. Each substance has a very characteristic absorp-
tion spectrum made up of a large number of absorption lines and
bands. Although the number of these absorption bands is
always in the thousands they can all be arranged in groups and
series in which the distribution may be formulated mathe-
matically and it is now possible to understand the physical
significance of this mathematical formulation.

According to the Bohr theory the absorption of light of definite
frequency takes place whenever the internal energy of the mole-
cule increases in the form of jumps. In the case of molecules,

three kinds of motion must be considered: (1) the orbital motion of the electrons, (2) the vibration of the atoms or groups of atoms, (3) the rotation of the molecule. From this point of view the internal energy $W°$ of a molecule in the normal state may be expressed as the sum of three quantities:

$$W° = E_e° + E_v° + E_r°$$

where $E_e°$, $E_v°$ and $E_r°$ are the values of the electronic, the vibrational and the rotational energies. When the molecule is activated under the influence of light the energy of the activated molecule will be

$$W' = E_e' + E_v' + E_r'$$

and according to the second postulate of Bohr the frequency v of the absorbed light is

$$\nu = \frac{W' - W°}{\cdot}$$

where h is Planck's constant and equal to $6.55.10^{-27}$ erg. sec.

If only the rotational energy of the molecule is considered it can be shown that this is dependent on two factors,—the quantum number of rotation and the moments of inertia of the molecule. If the molecule is isotropic there is only one moment of inertia J. The successive states of rotation of the molecule correspond to successive values of the internal energy $E_r°$, E_r', E_r'', . . . E^m, E_r^{m+1} . . . the general formula for a molecule with one moment of inertia being

$$E_r^m = \frac{h^2 m(m + 1)}{8\pi^2 J}$$

The transition from one state m to the next $(m + 1)$ or to the preceding $(m - 1)$ state absorbs light of definite frequency. The general distribution of these absorption lines is given by the formula

$$\nu = \nu_0 + \frac{E_r^{m+1} - E_r^m}{h} \quad \text{or} \quad \nu = \nu_0 + \frac{E_r^m - E_r^{m-1}}{h}$$

This distribution depends upon only *one* value which is characteristic of a particular molecule, that is the moment of inertia. This means that in all parts of the absorption spectrum, the fine structure, which is determined by the rotation of the molecule, must be the same. The rotational spectrum of a molecule with one moment of inertia shows a series of equidistant lines in which the distance between two successive lines is equal to

$$\Delta \frac{1}{\lambda} = \frac{h}{4\pi^2 J} = \frac{55.5 \cdot 10^{-40}}{J} \text{ cm}^{-1}$$

If therefore the distance between successive lines is measured the moment of inertia can be readily calculated. This conclusion has been confirmed by a great number of measurements on different molecules with only one moment of inertia,—nitrogen, hydrogen, sulphur, hydrogen chloride, carbon monoxide, etc.

In the case of the molecule of methane the measurements made by Cooley (8) in the infra-red have shown that there are two different sets of fine structure. For the band $\lambda = 3.3\mu$ there are many narrow lines in which $\Delta\frac{1}{\lambda} = 5.51$ cm^{-1}, but for the band $\lambda = 7.7\mu$ the lines are more widely spaced and $\Delta\frac{1}{\lambda} = 9.77$ cm^{-1}. This structure of the absorption spectrum cannot be that of a molecule with only one moment of inertia. During the last year these results have been discussed quite fully by Dennison (9) and particularly by Guillemin (10) from the laboratory of Sommerfeld in Munich.

The absorption spectrum of formaldehyde has been studied in my laboratory with Dr. Schou (11) and we have found a double rotational spectrum with two sorts of fine structure. The particular importance of these results depends upon the fact that if the molecule has two moments of inertia, J and K, it is possible to calculate these values from the distribution of the fine lines in the absorption spectrum.

When a molecule has two moments of inertia, J and K, the energy of rotation depends on two quantum numbers m and q. The energy of rotation in any state E_r^{mq} is expressed by

$$E_r^{mq} = \frac{h^2}{8\pi^2}\left[\frac{m(m+1)}{K} + q^2\left(\frac{1}{J} - \frac{1}{K}\right)\right]$$

in which q is constant in the transition from the state m to (m $\pm$ 1) and m remains constant in the transition from q to (q $\pm$ 1), thus giving rise to two sorts of absorption lines. By application of this formula the values for the two moments of inertia may be calculated. For the methane molecule J = $3.64 \cdot 10^{-40}$ and K = $5.65 \cdot 10^{-40}$ and for the formaldehyde molecule J = $1.41 \cdot 10^{-40}$ and K = $25 \cdot 10^{-40}$. These results show that the molecule of methane cannot be a tetrahedron and the simplest structure which is in accord with these values is that of a pyramid. From the values of the two moments of inertia, corresponding to the rotation around the two axes, it is possible to calculate the distance between the different atoms. Guillemin's calculations (10) show that the distance between the hydrogen atoms in methane is $1.05 \cdot 10^{-8}$ cm. and the distance between the carbon and hydrogen atoms is $1.15 \cdot 10^{-8}$ cm., the height of the pyramid being equal to $0.375 \cdot 10^{-8}$ cm. For formaldehyde we have found the distance between the carbon and oxygen atoms to be $1.02 \cdot 10^{-8}$ cm. and between the two hydrogen atoms $1.34 \cdot 10^{-8}$ cm.

3. THE STRUCTURE OF CRYSTALS AS DETERMINED BY X-RAYS

It is not possible to examine the crystal structure of solid methane but different derivatives of methane have been examined. The crystal structure of pentaerythritol $(C(CH_2OH)_4)$ has been studied by Mark (12) and his results have been confirmed by Huggins and Hendricks (13). The results show that this compound cannot have cubic symmetry and that the most probable structure of the molecule which is in agreement with the experimental values is that of a pyramid. Similar results have been obtained with tetraphenyl methane and on the basis of these results Weissenberg (14) and Reis (15) have developed a whole new stereochemistry based on the conception of the pyramidal structure of the methane molecule.

4. POTENTIAL ENERGY OF THE MOLECULE OF METHANE

The theoretical study of the stability of different configurations of a molecule formed by one negative ion and one or more hydrogen ions has been investigated recently by Heisenberg (16), Born (17), Kornfeld (18), Hund (19), and the molecule of methane in particular by Guillemin (10). This theoretical study is based upon the following considerations: (1) between the carbon atom, or rather the carbon ion, and each hydrogen ion there is first an attractive force proportional to $1/r^2$ and secondly, a repulsive force proportional to $1/r^n$, where n = 5 or 7 or 9; (2) between the different hydrogen atoms there are forces propor-
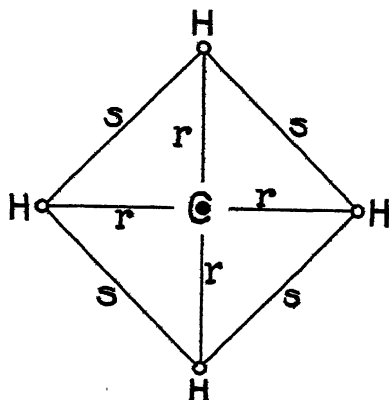
Fig. 9

tional to $1/s^2$ and to $1/s^n$; (3) inasmuch as the carbon ion is polarized or deformed by the intramolecular electrical field, the resulting dipole exerts an additional action on the hydrogen ions and this action is dependent upon the deformation coefficient α of the carbon ion ($\alpha = 1.34 \cdot 10^{-24}$). The potential energy of a configuration such as is shown in figure 9 can be expressed as the sum of terms depending upon $1/r$, $1/s$, and α:

$$P = F(1/r) + G(1/s) + H(\alpha, 1/r)$$

This potential energy has a minimum value for a certain definite configuration and the conditions for this minimum value can be

calculated. The form corresponding to the minimum of the potential energy is dependent upon the value of the deformability coefficient α. The results of these calculations show that the potential energy of the tetrahedral configuration is larger than that of the pyramidal form since for the tetrahedral form $P = -136 \cdot 10^{-12}$ erg and for the pyramidal $P = -175 \cdot 10^{-12}$ erg.

These results show therefore that the stable form of the methane molecule is that of a pyramid and not of a tetrahedron. The same method of calculation when applied to the structure of the water molecule shows a triangular structure (fig. 10) with an angle of 66° and to the molecule of ammonia a pyramidal form (fig. 11).

Another important result derived from the mathematical treatment of potential energy is obtained when this method of calcula-
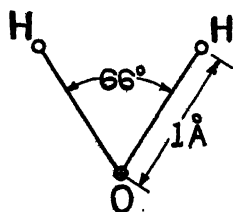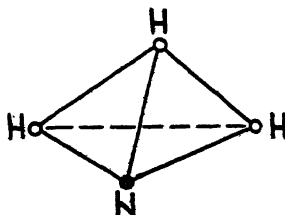
FIG. 10   FIG. 11

tion is applied to the molecule of carbon tetrachloride. In this case the structure of minimum potential energy is the tetrahedral form and not the pyramidal. These conclusions are of particular significance inasmuch as they show not only that the structure of the methane molecule is pyramidal but also that the form of a molecule cannot be considered as rigid and fixed. It is incorrect to conclude that because a molecule has a certain definite structure that this same type of structure would be maintained in all derivatives. On the contrary, a molecule is a mobile system of atoms in which the stability of the structure is determined by the minimum value of the potential energy. Instead of assuming the *rigidity* of the molecule, which has been the assumption upon which organic chemistry has been developed in the past, there must be a recognition of the *lability* of every molecule. This

lability of structure determines the chemical reactivity and if the values of the potential energy for different molecular structures are known it is possible to predict the chemical reactivity.

5. EMISSION SPECTRA OF THE CARBON ATOM

The carbon atom contains two inner K-electrons which have circular orbits and correspond to the quantum notation 1_1, and four external or valence electrons. Until 1922 it was assumed that these external electrons were all of the same type, with circular or 2_2 orbits. In 1924, however, Fowler (20) in studying the emission spectra of ionized carbon (C+) found that this spectrum had the same structure as that of boron. Analysis of the boron spectrum had already shown that the boron atom contained two external electrons with 2_2 orbits and one electron with a 2_1 orbit. It would follow therefore that the ionized carbon (C+) must have the same two types of external electrons. The structure of the carbon spectrum for the normal carbon atom (C) and the successive ions, C+, C++, C+++, and C++++ has been studied recently by Millikan and Bowen (21) and the general result of their work shows that there must be two types of valency electrons in the carbon atom. Two of these electrons have circular or 2_2 orbits and two have elliptical or 2_1 orbits. It would follow therefore that the four valencies of carbon would not be equivalent and there must be a distinction made between the two types of valencies. Further analysis of the emission spectra of carbon gives a method of determining values of the energy required to bring about the successive ionizations of the carbon atom. The results are as follows:

$$C \rightarrow C_+ + e - 184{,}000 \text{ calories}$$

$$C_+ \rightarrow C_{++} + e - 560{,}000 \text{ calories}$$

$$C_{++} \rightarrow C_{+++} + e - 1{,}060{,}000 \text{ calories}$$

$$C_{+++} \rightarrow C_{++++} + e - 2{,}300{,}000 \text{ calories}$$
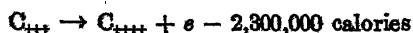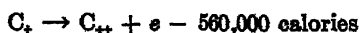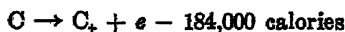
SUMMARY

By the application of the modern methods of molecular physics to the study of the structure of the methane molecule, the following conclusions, based on five completely independent methods have been established:

1. The four valencies of carbon are not equivalent but consist of two different types.

2. The structure of the molecule of methane is pyramidal in form and not tetrahedral as previously assumed.

3. The molecule of methane is a labile system of atoms and is capable of change to other structures in its derivatives.

REFERENCES

(1) TYNDALL, J.: Phil. Mag. (4) 37, 384 (1869).

(2) STRUTT, J. W.: Phil. Mag. (4) 41, 107, 274, 447 (1871).

(3) STRUTT, R. J.: Proc. Roy. Soc. 95A, 155 (1918); 97A, 435 (1919); 98A, 57 (1920); 102A, 190 (1922); Compt. rend. 171, 1002 (1920).

(4) CABANNES, J.: Compt. rend. 171, 852 (1920); J. phys. radium 1, 129 (1920).
    CABANNES, J., AND LEPAPE, A.: Compt. rend. 179, 325 (1924).
    CABANNES, J., AND GRANIER, J.: J. phys. radium 4, 429 (1923).

(5) GANS, R.: Ann. Physik. 65, 97 (1921).

(6) RAMAN, C. V.: Nature 109, 75 (1922); 114, 49 (1924).
    RAMAN, C. V., AND RAMANATHAN, K. R.: Phil. Mag. 45, 113 (1923).
    RAMAN, C. V., AND RAO, K. S.: Phil. Mag. 46, 426 (1923).

(7) CABANNES, J., AND GAUZIT, J.: J. phys. radium 6, 182 (1925).
    CABANNES, J., AND GRANIER, J.: Compt. rend. 182, 885 (1926).

(8) COOLEY, J. P.: Astrophys. J. 62, 73 (1925).

(9) DENNISON, D. M.: Astrophys. J. 62, 84 (1925).

(10) GUILLEMIN, V.: Ann. Physik. 81, 173 (1926).

(11) HENRI, V., AND SCHOU, S. A.: Compt. rend. 182, 1612 (1926).

(12) MARK, H., AND WEISSENBERG, K.: Z. Physik. 17, 301 (1923).

(13) HUGGINS, M. L., AND HENDRICKS, S. B.: J. Am. Chem. Soc. 48, 1647 (1926).

(14) WEISSENBERG, K.: Ber. 59B, 1526 (1926).

(15) REIS, A.: Ber. 59B, 1543 (1926).

(16) HEISENBERG, W.: Z. Physik. 26, 196 (1924).

(17) BORN, M., AND HEISENBERG, W.: Z. Physik. 23, 388 (1924).

(18) KORNFELD, H.: Z. Physik. 26, 205 (1924).

(19) HUND, F.: Z. Physik. 31, 81 (1925).

(20) FOWLER, A.: Proc. Roy. Soc. 105A, 299 (1924).

(21) BOWEN, I. S., AND MILLIKAN, R. A.: Phys. Rev. 25, 594 (1925); 26, 150 (1925); 26, 310 (1925); Proc. Nat. Acad. 10, 199 (1924); Nature, 114, 380 (1924).

■

PROGRESS IN THE STRUCTURAL STUDY OF CARBOHYDRATES[1]

JAMES COLQUHOUN IRVINE

University of St. Andrews, Scotland

My first words must be of the debt I owe to my fellow-chemists in America who have paid me the singular compliment of adding my name to the list of Willard Gibbs Medallists. It is the custom, I know, to mark the presentation of the medal by the delivery of an address on the researches for which the award is made, and custom cannot be defied; but if for a moment I may be permitted to speak of my feelings rather than of my work, it would be to say that this honour has never been bestowed on one who prized it more than I do. To me, the distinction carries a double significance, for I regard it as an outward symbol of the personal friendship extended to me by the chemists of America, a friendship which I count as not the least of fortune's favours, and I esteem this award in even higher measure through its association with the name of Willard Gibbs.

It would be ungracious of me, also, if I failed to take this opportunity of expressing the gratitude I have always felt to the scientists of the New World for another gift, one which came long ago, yet gradually and almost imperceptibly. I refer to the encouragement, denied alas to so many who are worthy of it, which comes to a scientific worker when he finds, as I found in this country, that his work is closely studied, and that, having been weighed in the balance, it is not found wanting. I can never forget the stimulus, the exaltation, derived from your understanding and sympathetic interest in the problems with which I am engaged. The lot of the investigator is often made lonely through lack of a word of encouragement, but, throughout

[1] An address delivered on the occasion of the Willard Gibbs Medal Award, Chicago, Illinois, September 17, 1926.

life, it has been my good fortune to be surrounded by indulgent friends.

I turn now to my address, finding comfort in the fact that organic chemists are deemed worthy of the Willard Gibbs Medal, and that some of my predecessors share with me the misfortune that their activities have lain in fields remote from that vast country opened out by the instinctive genius of the man whose memory we honour to-night. It is, indeed, a far cry from the Structure of Carbohydrates to the 'Phase Rule, yet although there may be no very obvious connexion between them, a relationship is nevertheless there, for the work I hope to summarise for you could not have been contemplated, far less accomplished, without daily application of the principles laid down by the physical philosopher of New England. The year in which the classical papers of Willard Gibbs appeared was the year of my birth; his views had reached the lecture room before I vacated the students' bench and, throughout my reseach life, I have been under the influence he radiated on British science. It is a debt of which I am fully conscious, and I add my tribute to his inspiration.

This is the 17th of September, 1926, and memory warns me that, exactly a quarter of a century ago to a day, I commenced the first of what has proved to be a continuous series of investigations on the chemistry of sugars. In the intervening years I have often been asked why my attention was drawn to this particular class of compound, and this is a question which might well be put to many others, for it must be evident, even to the casual student of organic chemistry, that there has been in recent years a marked expansion in the study of carbohydrates. Whereas at the beginning of the century there were perhaps a dozen expert workers in this field, their successors are now legion, and the number grows almost daily. It has been an amazing development, and one for which it is difficult to find an adequate explanation, but, looking back on my own experience, I think I can discern the underlying reasons. Certainly among these reasons, simplicity can find no place, either in respect of

manipulation or of speculation. Even in the investigation of the compounds which, with unconscious irony, we term the "simple" sugars, the organic chemist is confronted with formidable experimental difficulties, to be appreciated fully only by those who have submitted themselves to this stern discipline. Months may be spent in attempts to induce a syrup to crystallise; and, even if this be successfully accomplished, the product may prove to be one of twenty distinct but closely-related isomeric forms. Nor can the prospect of material reward play any considerable part in attracting workers to the chemistry of sugars, for, so far at least, our sugar factories and even our cellulose industries depend only in minor degree on the power to elucidate molecular constitution. We must therefore look for other causes, specific to the chemistry of sugars, which invest these compounds with a special fascination. My reflections have convinced me that first importance must be attached to the fact that carbohydrates are essentially the products of natural, as opposed to artificial, synthesis. Within us and around us, through the agency of life, these compounds are being built up and broken down, formed and transformed, finally to be consumed in the fire of metabolism. To study the sugars is, in brief, to study the great molecular channel through which solar energy flows to us. No wonder the thought of such work makes a powerful appeal to the student of vitalism who is naturally somewhat inclined to turn aside from that aspect of chemistry which deals with purely artificial reactions. Then again, one must recognise that to be an intelligent investigator of the sugars the chemist must be something of a biologist and of a physiologist, but, above all, he must be a physicist. This diversity of interest finds its expression, not only in the library but also in the laboratory where one day our chemist may be manipulating many kilograms of plant products only to turn later to the other end of the working scale to examine a few milligrams of a precious crystalline sugar. He has to be familiar with bacteriological, as well as with orthodox chemical processes; he finds himself engulfed in the mysteries of colloids and, at all times, the precision instruments of the physi-

cist are his daily tools. Taking a broad and, I trust, an impartial view of these requirements, the marvel to me is not so much that many investigators turn to the sugar group, but that so many can resist its attractions.

There remains another factor, as potent, I believe, as any I have mentioned, and that is expressed in the statement that to-day the sugars are being formulated in terms of molecular structure, so that at last we are approaching the position when we may study their reactions with precision and understanding.

There is a special satisfaction in discerning how our knowledge of carbohydrates has expanded in response to the stimulus conveyed by fresh views on structure. So long as glucose was regarded merely as a molecule $C_6H_{12}O_6$, characterised mainly through its power to reduce Fehling's solution and to rotate the plane of polarised light to the right, the rôle of the compound in natural processes, of necessity, remained obscure. An entirely new situation was created when we began to recognise that within the glucose molecule there are five hydroxyl groups, and a further step was taken when it became apparent that these groups varied in character and property.

GLUCOSE

This brings me to a point where I may indicate the general position which had been reached in 1901 when, as I have indicated, my share in this work commenced. The first phase of Fischer's classical researches was over; the most important monosaccharides had been synthesised, and the principles of stereoisomerism had been made the basis of a rational classification which accounted for all the hexoses and pentoses then known. Nevertheless the chemistry of sugars was then in a state of flux. Limiting our attention meanwhile to the fundamental case of glucose, a sugar which I believe to be the origin of all carbohydrates, the old aldehydic formula for the compound had been replaced by a cyclic structure in which an oxygen ring was represented as connecting the first and fourth carbon atoms as shown below:

```
CHO              1--    -------  ┌──CHOH
 │                                │
CHOH             2--    -------│  CHOH
 │                          O  │
CHOH             3·------------·│  CHOH
 │                                │
CHOH             4------------ └──CH
 │                                │
CHOH             5-----·--------CHOH
 │                                │
CH₂OH            6--------------CH₂OH
```

Aldehydic or Straight Chain formula for glucose	Butylene-oxide formula for glucose
I	II
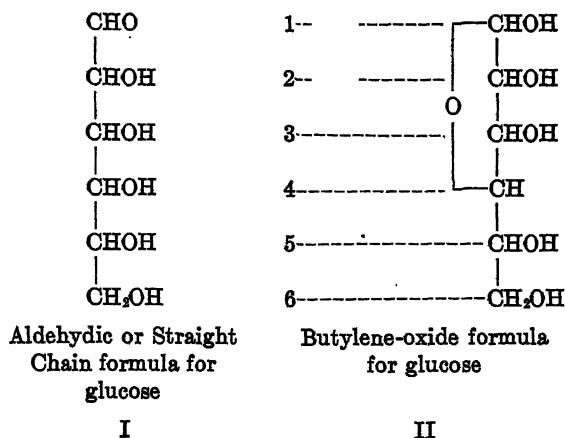
I need not enumerate here the reasons which led to the opinion that the reducing sugars should be formulated as cyclic compounds, but may perhaps state that the adoption of the five-membered ring was dictated largely by analogy, rather than by direct experimental proof. Nevertheless, the adoption of such a formula marked a substantial advance. For example, it was thereafter possible to recognise that the hydroxyl groups of a reducing sugar such as glucose may be classified into three categories: (a) The terminal hydroxyl group attached to carbon atom No. 1, which, being readily oxidised, is responsible for the reducing action on Fehling's solution. This group is feebly acidic in character, and is not only highly reactive, but is also capable of transposition giving rise to the α- and β-forms of the sugar. (b) The hydroxyl groups within the ring. These are characterised by less abnormal properties and remain rigidly fixed in position, thereby accounting for the stereochemical distinction between the isomeric aldoses. (c) The hydroxyl groups external to the ring, the most important being the primary alcohol group, which takes part in the oxidation reactions leading to the formation of such compounds as the glycuronic acids.

I have given only an indication of the advantages secured through classification of the hydroxyl groups in glucose, but possibly I have said sufficient to indicate that the complete

exploration of a sugar molecule demands the study of the individual asymmetric systems, for it must be remembered that each contributes its own specific properties to those of the molecule as a whole.

Evidently twenty-five years ago the time was opportune to engage in this highly specialised study of sugar structure, but even more attractive problems were presented by the prospect of opening out a path whereby the molecular constitution of disaccharides and polysaccharides could be approached. I need not remind you that the disaccharides may be regarded as compound sugars formed by the union of two hexoses through loss of a molecule of water. When it is considered that some disaccharides are devoid of action upon Fehling's solution, it is at once clear that the reducing groups of each of the constituent molecules must have taken part in the formation of the complex sugar. Similarly, in the case of reducing disaccharides, it is evident that the reducing groups of only one of the constituent glucose molecules is involved in the union, whilst the other remains free. This much was known, but, beyond that, all ideas of the exact positions through which one sugar united with another to form a disaccharide were purely speculative. The great master Emil Fischer clearly recognised this limitation imposed by the lack of suitable experimental processes, and abstained from anything more than tentative suggestions as to the molecular structure of the disaccharides. Would that his caution had descended on his successors.

With regard to polysaccharides, the situation at the beginning of the century was even more obscure. Within the single empirical formula $(C_6H_{10}O_5)_n$ were accommodated such diverse compounds as cellulose, starch and glycogen, but beyond the knowledge that, as a purely arithmetical ratio, three hydroxyl groups were present for each six carbon atoms, remarkably little evidence was available which could be brought to bear on the problem of molecular constitution. To sum up, the position was that structural sugar chemistry ended with the monosaccharides.

METHYLATED SUGARS

I have outlined some of the carbohydrate problems which then confronted the chemist, and it is my belief that the most effective method of solving these structural mysteries has been found in the study of methylated sugars. It was with this hope in my mind that I commenced my work. A word of explanation is perhaps necessary as to what is meant by the expression "methylated sugar." Clearly, if the sugars are poly-alcohols, they should be capable of forming the corresponding ethers, and these should display the stability, particularly towards hydrolysts, which is so characteristic of the ether type. A first step in this direction had been taken by Emil Fischer who discovered the general reaction of converting reducing sugars into the corresponding methylglucosides. In such compounds, however, the methyl group is present in the terminal position, and is readily lost on hydrolysis. The alkyl glucosides are therefore more closely akin to esters than to ethers. It was evident that, if normal poly-ethers of the sugars could be obtained, the key would be found to decipher the complete structure of the complex sugars. Let us take an example. In maltose, we have two glucose residues attached by the reducing group of one of them to a non-reducing group of the other. There are, however, four such groups available and the question is, which of them is involved in the coupling? It is the case that the hydroxyl positions in the disaccharide may be substituted by acetyl or other acyl groups, and that the point of junction between the constituent hexose residues cannot take part in such substitution. But, if we disrupt the acetylated disaccharide by means of hydrolysis, we eliminate at the same time the substituting acyl groups, and thus all structural evidence is lost. An entirely different state of affairs was to be expected in the case of a methylated disaccharide, as the methyl groups would survive hydrolysis and hydroxyl groups would be regenerated in the positions through which the constituent sugars had been joined together. Thus, in the case of maltose, for example, we should obtain as the hydrolytic products a tetramethyl glucose and a

trimethyl glucose.   In one of these methylated sugars only a
reducing hydroxyl group can be present, (Let us call it A), whilst
in the other there must be, in addition, one non-reducing hydroxyl
group (Let us call it B).   Obviously, we would then be in the
position of regarding maltose as a compound sugar consisting
of two glucose residues condensed together through the positions
we have indexed as A and B.   The principles I have enunciated
are capable of general application, and can be extended to the
constitution of glucosides and polysaccharides, but the exploita-
tion of the idea demanded that much preliminary research had to
be undertaken.   The first consideration was to find a method of
alkylation suitable to the sugar group, and the other was the
preparation of a sufficiently large variety of fully and partly
methylated sugars which would function as reference compounds
in extending the general research scheme.   Fortunately, a con-
venient experimental process of methylation lay in my hands.
You will remember the elegant method discovered by Purdie, my
distinguished predecessor in the Chair of Chemistry at St.
Andrews, by means of which remarkably smooth alkylation is
effected by submitting hydroxy-compounds to the joint action
of methyl iodide and dry silver oxide.   In applying this process
to the particular case of reducing sugars complications were in-
troduced through the oxidising effect of the silver oxide, but these
were overcome by the use of methylglucoside as the starting
material.   This compound was successfully converted into a
tetramethyl methylglucoside which, following the general rule,
lost only the glucosidic group on hydrolysis and, accordingly,
yielded a tetramethyl glucose as the final product.   This was
the first methylated sugar obtained and, as originally formulated,
the compound was represented as a butylene-oxide.   In the
meantime, this formula may be retained, although the general
question of the different positions assumed by the oxygen-ring
in glucose must engage our attention at a later stage.   Tetra-
methyl glucose represents only one type of a methylated sugar,
but other types containing a smaller number of alkyloxy-groups
are known.   Of these, we may select as an example, one partic-

ular variety of monomethyl glucose. Commencing with glucose-diacetone, a compound in which four hydroxyl groups of the sugar are already substituted, methylation can give, as a maximum, only a monomethyl derivative. Subsequent hydrolysis removes the unstable acetone residues, but leaves the methyl group unaffected, and a monomethyl glucose results. This may be taken as representative of the general method employed to prepare other partly methylated sugars of which numerous examples are now known. Some of the properties of methylated sugars may be mentioned at this stage as, in many respects, they differ profoundly from those of the parent compounds. As is to be expected, the step-by-step replacement of hydroxyl by methoxyl results in a progressive increase in stability and solubility, so·that in the case of tetramethyl glucose, the sugar dissolves in all ordinary solvents, ranging from water to petroleum ether. The compound is also so stable that it can be distilled quantitatively in a vacuum.

Characterised as they are by such convenient properties, methylated sugars have proved of great value in the detailed study of such problems as mutarotation and the determination of the molecular rotations of the α- and β-forms of sugars. Further, their use enables the investigation of sugars to be confined to selected hydroxyl groups, thereby revealing the characteristic properties of each asymmetric system in the molecule. The fascination of methylated sugars is such that one might have continued to explore their properties to the exclusion of the ultimate object in view in preparing such compounds.

THE OXYGEN RING

This brings me to the results obtained in the application of methylated sugars to the structural problems of complex carbohydrates, but, as a preliminary to this discussion, let us glance for a moment at the question of the position occupied by the oxygen ring in the simple sugars. As I have already explained, the view originally held was that glucose is definitely a butylene-oxide.

```
    ┌─────CHOH
    │        │
    │       CHOH
    │        │
  O │       CHOH
    │        │
    └─────CH
             │
            CHOH
             │
            CH₂OH
```
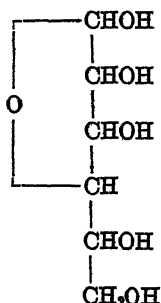
That the ring does not, of necessity, remain in one fixed position was revealed about twelve years ago, when Fischer isolated a form of methylglucoside which differed from the known α- and β-varieties. Following the general but regrettable custom, the new product was termed by him "γ-methylglucoside," and he. recorded that, when hydrolysed, it yielded ordinary glucose. Application of the methylation process at once gave the explanation as to why methylglucoside should exist in a larger number of forms than stereochemical considerations accommodate, as a new variety of tetramethylglucose utterly distinct from the crystalline sugar to which I have referred was thereby produced. Consideration will show that the only possible reason for the existence of these isomeric tetramethyl glucoses must lie in the different positions assumed by the oxygen ring in each compound. The introduction of the methyl groups prevents the ring from reverting to a stable from an unstable position, and thus methylated sugars are invested with increased importance, and acquire a new sphere of usefulness.

The expression "γ-sugar" crept into use to designate aldoses and ketoses in which the internal ring is displaced from the normal stable position, and has therefore been applied to an entirely new, highly reactive class of sugars. In view of these results, I would urge for extreme caution in accepting any suggestion that glucose is definitely an amylene-oxide. There can be no fixed and unalterable structure for glucose, as it is conceivable that the oxygen-ring may couple position 1 with any

of the five remaining positions, so that if we include an aldehydic variety, d-glucose may actually exist in eleven forms and may react as a mixture of any of these. It is not a case of one oxygen-ring, but the option of several alternatives and, in consequence, there can be no formula universally applicable to glucose. It follows also that the ring-structure of a sugar must be investigated separately in each derivative studied.

Examples are now available in which this has been done. Recently, it has been shown that, when tetramethyl glucose is oxidised, a trimethoxy-glutaric acid is found amongst the oxidation products. If this be accepted, the methylated sugar must be an amylene-oxide and the constitution of certain natural glucosides, such as salicin, arbutin and indican, must conform to the following structural type:

$$-\text{CH}- \qquad -\text{O}- \qquad -\text{R}$$

$$
\begin{array}{c}
\text{CHOH} \\
| \\
\text{O} \quad \text{CHOH} \\
| \\
\text{CHOH} \\
| \\
\llcorner\!\!-\text{CH} \\
| \\
\text{CH}_2\text{OH}
\end{array}
$$
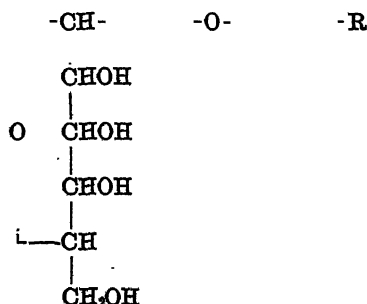
III

Beyond this it is inadvisable to go.

Although a considerable variety of methylated hexoses were prepared as a preliminary to the constitutional study of di- and polysaccharides, it has been ascertained in practice that only five of these find application. The sugars in question are: tetramethyl glucose, two isomeric forms of trimethyl glucose, tetramethyl γ-fructose and tetramethyl galactose. The compounds are formulated below, but it is perhaps necessary to emphasise that only in the cases of tetramethyl glucose and of tetramethyl galactose are we justified in allocating a fixed position to the internal oxygen ring.

```
 ┌────CHOH         ┌────CHOH         ┌────CHOH         CH₂OMe
 │    CHOMe        │    CHOMe        │    CHOMe     ┌────COH
 ⌐ CHOMe         O  CHOMe         O  CHOMe         │    CHOMe
 │    CHOMe        │    CHOH         │    CHOMe     ·O   CHOMe
 └────CH          └────CH          └────CH          └────CH
     CH₂OMe           CH₂OMe           CH₂OH            CH₂OMe
```
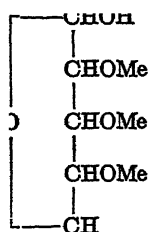
2:3:4:6-Tetra- methyl glucose or galactose.	2:3:6-Trimethyl glucose.	2:3:4-Trimethyl glucose.	Tetramethyl γ-fructose.
IV	V	VI	VII
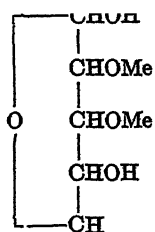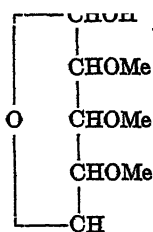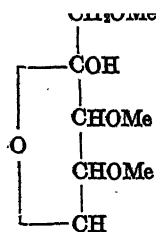
Greater complications are encountered when we consider the two trimethyl glucoses. In one of these sugars (2:3:4-trimethyl glucose), which is a colourless liquid resembling glycerol in appearance, the unsubstituted hydroxyl group is definitely in the No. 6 position, and the formation of this compound from a disaccharide is proof that the one hexose residue is attached to the terminal position of the other. The case presented by the isomeric 2:3:6-trimethyl glucose is more complex, as the oxygen ring undoubtedly fluctuates between the positions 1:4 and 1:5, thereby giving us the two alternative compounds represented by the following formulae:
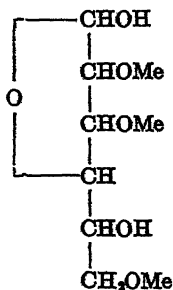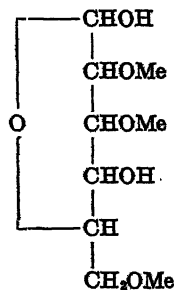
```
       ┌────CHOH              ┌────CHOH
       │    CHOMe             │    CHOMe
     O │    CHOMe           O │    CHOMe
       │    CHOMe             │    CHOH
       └────CH               └────CH
            CHOH                   CH₂OMe
            CH₂OMe
           VIII                    IX
```

Notice that this disconcerting property at once introduces an ambiguity into structural studies, in that the formation of this sugar from a disaccharide is not discriminative, as the coupling of

the constituents may be through either position 4 or position 5. It is necessary to make this point clear as, in the structural classification of disaccharides built up by Haworth, this important issue has been entirely overlooked. In addition, it is now recognised that his work includes a fundamental experimental error, which renders his whole scheme valueless, and that, consequently, so far as the constitution of disaccharides is concerned we are only at the beginning of things in place of at the end.

THE DISACCHARIDES

·Let us turn to the authentic results obtained by applying the methylation process to disaccharides. The first point to notice is that two of the diglucoses are found to behave in exactly the same way as shown below:

Maltose		2:3:4:6-Tetramethyl
and	on methylation	glucose.
Cellobiose	and hydrolysis →	and
		2:3:6-Trimethyl glucose.
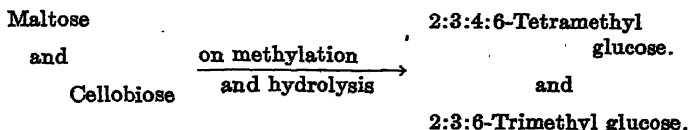
On first inspection this result might be interpreted as meaning that in maltose and in cellobiose the glucose residues are attached through the same positions, the one sugar being α-glucose-glucoside, and the other β-glucose-glucoside. Such is not the case, however, and it has to be remembered that we must also accommodate iso-maltose and iso-cellobiose within our structural scheme. It is impossible to do so unless we take into account the fact that 2:3:6-trimethyl glucose could arise from either of the following diglucoses:

Glucose 5-glucoside.

X

```
   —CHOH
      |
    CHOH
      |
 O   CHOH
      |
    CH——O——CH·CHOH·CHOH·CHOH·CH·CH₂OH
      |        L——————O——————J
   —CH
      |
    CH₂OH
```
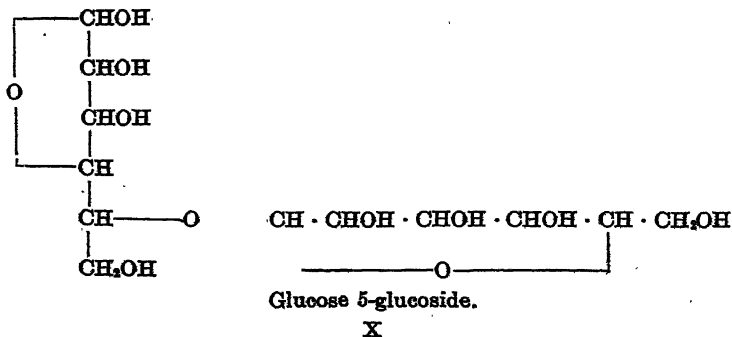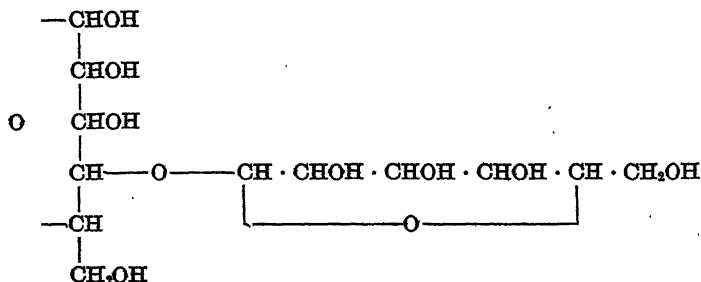
Glucose 4-glucoside.

XI

We are thus forced to the conclusion that one of these formulae represents the pair maltose and *iso*-maltose, whilst the other is reserved for cellobiose and *iso*-cellobiose. The same ambiguity is attached to the case of lactose, as the compound yields 2:3:4:6-tetramethyl galactose together with 2:3:6-trimethyl glucose. In consequence, if we regard the non-reducing section of the formulae given above as representing galactose, this residue may be attached either to position 4 or 5 of the glucose component. The situation, apparently uncertain, is greatly simplified when we remember that good reasons exist for the belief that starch and glycogen are composed of γ-glucose residues. This demands that maltose must be represented by Formula X in which case we are in a position to allocate Formula XI to cellobiose.

To the mass of evidence already available in support of this view, may be added the results recently contributed by Zemplén who, having degraded cellobiose to a gluco-erythrose, found that the product failed to form an osazone. This observation is consistent with the allocation of a glucose 4-glucoside structure to cellobiose as shown in Formula XI.

A clearer issue is discernible when we consider the case of gentiobiose, as this sugar, on methylation and hydrolysis, behaves as synopsised below:

2:3:4-Trimethyl glucose.

Gentiobiose⟨

2:3:4:6-Tetramethyl glucose.

In this case, there is no reason to doubt the fact that the hydrolytic products have been correctly identified and, further, the result is free from the complication of a dual interpretation, so that the structure of the parent disaccharide becomes:

$$
\begin{array}{l}
\text{-CHOH} \\
\;\;| \\
\text{CHOH} \\
\;\;| \\
\text{O}\quad\text{CHOH} \\
\;\;\;\;\;\;| \\
\text{CHOH} \\
\;\;| \\
\text{-CH} \\
\;\;| \\
\text{CH}_2\text{——O——CH} \cdot \text{CHOH} \cdot \text{CHOH} \cdot \text{CHOH} \cdot \text{CH} \cdot \text{CH}_2\text{OH} \\
\end{array}
$$

XII
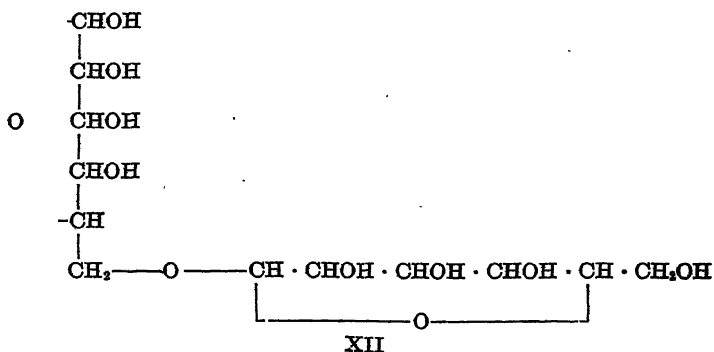
It is particularly gratifying to add here that one of the most recent results obtained by the American school of carbohydrate chemists has proved that the synthetic diglucose prepared by Fischer and termed by him "iso-maltose" is identical with gentiobiose, and thus possesses the above structure. This observation is important, as it will have the effect of removing another confusion from sugar chemistry through the deletion of the name "iso-maltose" for a disaccharide which is in no way related to the *iso*-maltose obtainable from starch.

Taking into account much supplementary evidence, a review of the whole situation shows that the reducing disaccharides based on glucose can be relegated to three structural types:

$$
\begin{array}{l}
\text{-CHOH} \\
\;\;| \\
\text{CHOH} \\
\;\;| \\
\text{O}\quad\text{CHOH}\qquad\qquad\text{-O-} \\
\;\;\;\;\;\;| \\
\text{CH——}\;\text{-O——CH} \cdot (\text{CHOH})_3 \cdot \text{CH} \cdot \text{CH}_2\text{OH}^* \\
\;\;| \\
\text{-CH} \\
\;\;\;\;| \\
\text{CH}_2\text{OH}
\end{array}
$$

XIII
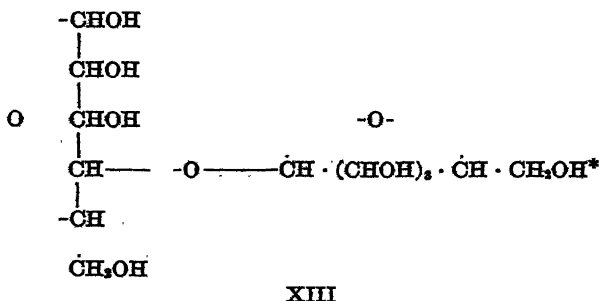
A represents cellobiose and *iso*-cellobiose, and
where * = galactose, also formulates lactose

$$\begin{array}{c} —\text{CHOH} \\ | \\ \text{CHOH} \\ \text{O} \quad | \\ \text{CHOH} \end{array}$$

-CH
 |
 CH————O————CH · (CHOH)$_3$ · CH · CH$_2$OH*
 |
 CH$_2$OH
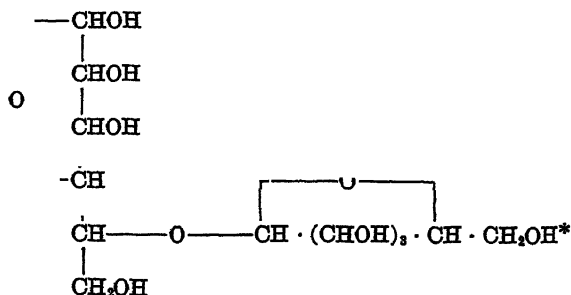
XIV

B represents maltose and *iso*-maltose, and where
* = galactose also formulates melibiose

C

—CHOH
 |
 CHOH
 |
O CHOH
 |
 CHOH
 |
——CH
 |
 CH$_2$————O————CH · (CHOH)$_3$ · CH · CH$_2$OH *
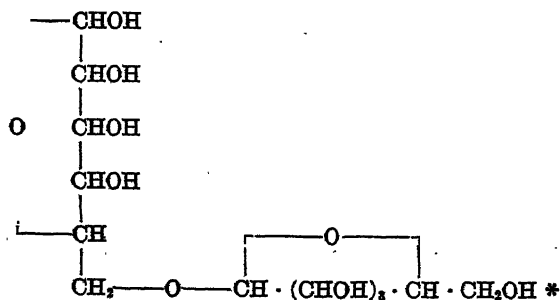
XV

C represents gentiobiose

So far I have left untouched the important case of sucrose, partly on account of the difficulty involved in clearing up a problem which, in itself sufficiently complicated, has been rendered more confused through the acceptance of results based on faulty work. It is probably the best known fact in sugar chemistry that sucrose may be hydrolysed to give dextro-rotatory glucose and laevorotatory fructose, but this has led to the idea that these constituents exist in the disaccharide in precisely the same form in which they are finally isolated. It is nevertheless the case that the fructose component of sucrose is not the ordinary laevorotatory variety of the sugar, but the unstable form

into which the sugars may pass through transposition of the oxygen ring. In other words, this component is γ-fructose.

This was the first serious theoretical complication encountered by Purdie and myself twenty-three years ago for, naturally, sucrose was the first disaccharide to which we turned our attention in developing the research scheme. An octamethyl sucrose was duly obtained and hydrolysed; the ordinary crystalline tetramethyl glucose was readily forthcoming as one of the scission products, but the fraction containing the tetramethyl fructose proved to be a liquid and, contrary to our expectation, was *dextro-rotatory*. This result can now be expressed in terms of the scheme

$$\text{Sucrose} \longrightarrow \text{Octamethyl sucrose} \begin{cases} \nearrow & \text{2:3:4:6-Tetramethyl glucose.} \\ \searrow & \text{Tetramethyl } \gamma\text{-fructose.} \end{cases}$$
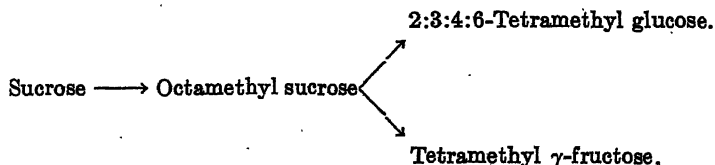
and the practical issue of the constitutional study of sucrose thereafter focussed on the structure of tetramethyl γ-fructose. Haworth, who investigated this problem, claimed with a confidence which from the first seemed unwarranted, that tetramethyl γ-fructose is an amylene-oxide and that, in consequence, sucrose should be represented by:

XVI
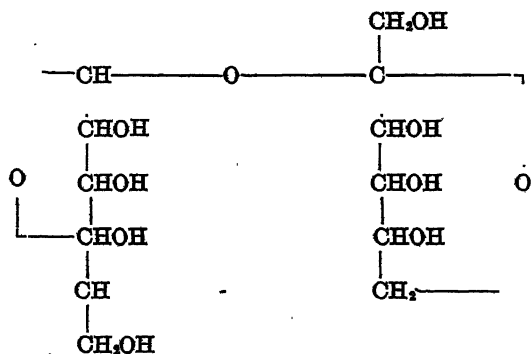
It is regrettable that the above formula has been widely adopted, as repetition of the work has failed to confirm the results which, it may be remarked, are further vitiated by a serious theoretical mistake. The oxidation of tetramethyl γ-fructose under accurately-controlled conditions has now revealed that the compound is either a propylene-oxide or a butylene-oxide, with much indirect evidence in favour of the latter view. It follows that the formula for sucrose which can be most strongly supported is:
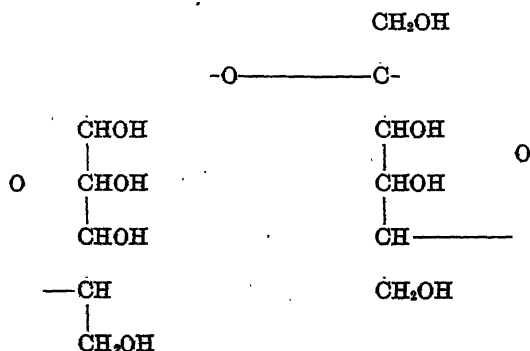
XVII

It may be remarked that part of the evidence leading to the above conclusion has been derived from studying those complex but interesting compounds formed by condensing glucose or fructose with acetone. This is another aspect of sugar chemistry which has been explored by the methylation process, and among the results obtained which are applicable to the problem now before us is the observation that whilst only γ-glucose condenses with acetone γ-fructose is incapable of this change which is restricted to the stable variety of the ketose.

The application of the methylation process as a means of determining structure can be extended far beyond the disaccharides, and has been used to examine the structure of the trisaccharide, raffinose. Unfortunately, on this question also, I must again disagree with both the evidence and the views submitted by Haworth. It is known that raffinose can be hydrolysed in two ways which result in the formation alternatively of sucrose or of

melibiose, and it follows that a satisfactory formula for the trisaccharide must incorporate the constitutions of these disaccharides. It was claimed that fully methylated raffinose gave on hydrolysis:

(a) 2:3:4:6-Tetramethyl galactose,

(b) 2:3:4-Trimethyl glucose,

(c) 1:3:4:6-Tetramethyl γ-fructose.

It is remarkable that this result is in agreement with formulae for melibiose and for sucrose which have been proved to be incorrect, and a scrutiny of the experimental details casts doubt on the identity of the trimethyl glucose actually isolated in the course of the work now referred to. The accepted formula for raffinose is thus incorrect in two out of the three hexose residues present, and the trisaccharide is much more probably represented by:

Galactose Residue.   Glucose Residue.   Fructose Residue.

Melibiose Residue.   Sucrose Residue.
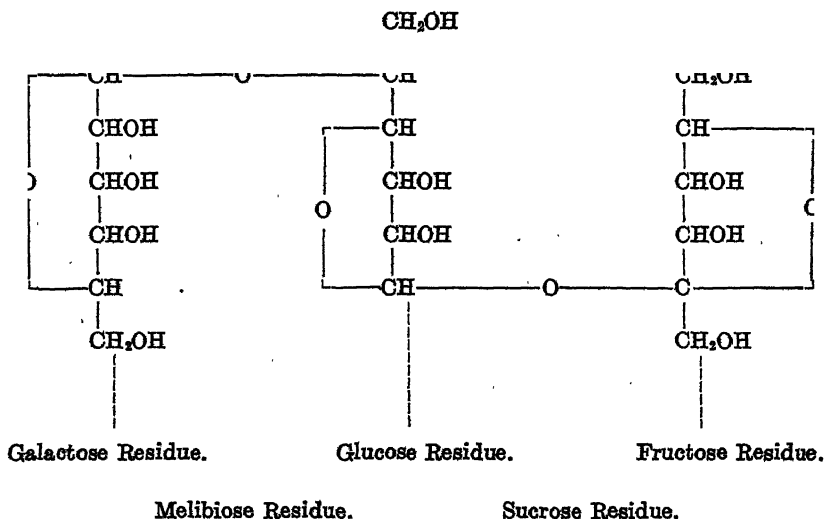
XVIII

There is, of course, one possible alternative, viz., that the attachment of galactose to glucose may be through position 4 in place of position 5 as represented above.

The desire to rectify the errors which have crept into disacch-

aride structure has forced me to spend a considerable time over these compounds, and we have now reached a stage when it is desirable to turn to another topic, and consider the progress in attacking the molecular structure of the polysaccharides. Speaking to this audience, I need not stress the point of the experimental obstacles which invested this section of our work with peculiar difficulty. These difficulties begin with the selection of the experimental material—if we ask ourselves the question, What are the criteria of standard specimens of cellulose or starch? we are left without an answer—and they multiply when it comes to isolating pure methylated derivatives from the colloidal systems in which they are formed. What I have called, perhaps somewhat slightingly, experimental obstacles must be regarded, however, as something more than mere inconveniences which have to be overcome, for it is only by close observation of every physical condition, as well as of every chemical property, that we can obtain a clear understanding of the polysaccharides. It would take too long to follow our various investigations in historical sequence, and I shall at once proceed to the results.

CELLULOSE

The first polysaccharide to be examined was cellulose, and it was quickly evident that the investigator, who approached this problem after experience with the subtleties of sugars proper, had to place himself in the position of a sceptical chemist. Two initial questions had to be faced: (1) is cellulose built up of glucose residues exclusively, and, (2) is cellulose based on large molecules constituted on the model of complex glucosides, or on simple anhydro-sugars highly polymerised. As a result of a series of strictly quantitative investigations, it was shown that cellulose may be converted quantitatively into glucose, and the way was thus ready to subject the polysaccharide to methylation. In this particular case, the silver oxide reaction was obviously inapplicable and, in consequence, we employed methyl sulphate and this methylation reagent has since become of general use. The introduction of the methyl groups proved to be a slow and

laborious operation, but ultimately a trimethyl cellulose was obtained as a friable mass which still preserved the external appearance and fibrous nature of the parent material. Hydrolysis gave 2:3:6-trimethyl glucose as the sole product, and at once it was revealed that cellulose is molecularly constituted on an exceedingly simple plan. The polysaccharide cannot be a complex glucoside of the type first advocated by Hess, but must originate by polymerisation of a simple anhydro-n-glucose. There remains the question of the magnitude of $n.$, and here we have, as a guide, the fact that cellulose is convertible into the disaccharide cellobiose. It follows that the simplest molecular unit for cellulose which would satisfy this condition is an anhydrocellobiose, and such a compound, in view of the quantitative conversion of cellulose into 2:3:6-trimethyl glucose may be formulated in only one way, provided we accept the structure of cellobiose already developed:

$$
\begin{array}{l}
\text{r---CH----------------------------O} \\
\quad\quad| \\
\quad\text{CHOH} \\
\quad\quad| \\
\text{O}\quad\text{CHOH} \\
\quad\quad| \\
\quad\text{CH--------O----------CH CHOH CHOH CH CH CH}_2\text{OH} \\
\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad| \\
\text{---CH}\quad\quad\quad\quad\quad\quad\text{L-------O-------} \\
\quad\quad| \\
\quad\text{CH}_2\text{OH}
\end{array}
$$
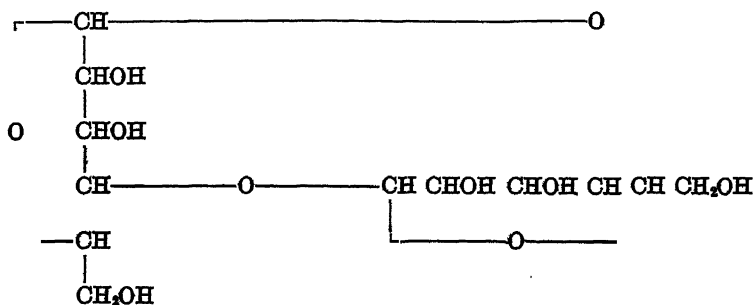
<div align="center">XIX</div>

It is to be remembered, however, that, even under the most favourable conditions, the yield of cellobiose obtainable from cellulose is far short of that demanded by the above formula, so that the next possibility to be considered is that the unit may be an anhydro-trisaccharide. We are now confronted with two possible formulae, as the anhydro-cellobiose postulated above may be expanded in two ways, so as to include a third glucose residue. In discriminating between these alternatives it is undesirable to speculate too freely but, so far as our present

knowledge goes, the most satisfactory formula for an anhydro-trisaccharide related to cellulose is:

$$
\begin{array}{l}
\text{CH}_2\text{OH} \\
| \\
\text{CH}\text{————————}\text{O} \\
| \qquad\qquad\qquad | \\
\text{CH CHOH CHOH CH}
\end{array}
$$
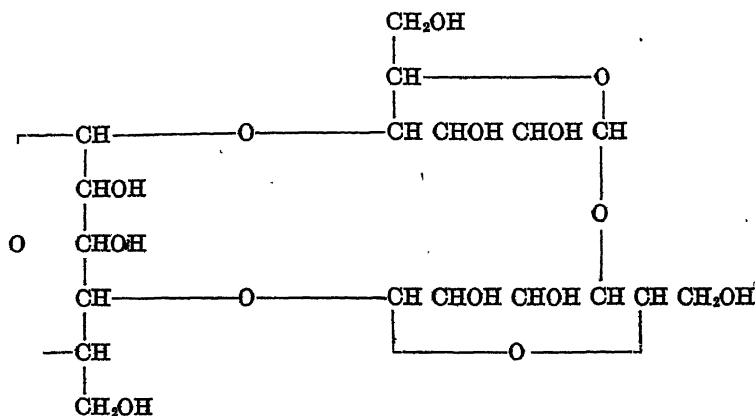
XX

Putting this idea to the test, cellulose has been depolymerised under accurate conditions, until the simplest possible dextrins are formed, and the result then obtained is striking. The product is naturally a mixture, but the essential constituent is the dextrin $(\text{C}_6\text{H}_{10}\text{O}_5)_3$. Here again, the methylation process has been valuable, as by its aid this anhydro-trisaccharide has been converted into soluble derivatives, by means of which it was possible to establish the all-important matter of the molecular weight.

It must not be assumed, however, that cellulose is exclusively built up of anhydro-triglucose molecules. The experimental evidence so far available justifies no more than the statement that at least one-third of cotton cellulose consists of polymerised anhydro-triglucose, while the results of x-ray spectrographic analysis indicates that the C_6 residues are marshalled in even numbers, probably in groups of four or six. Taking both types of evidence into account, my view is that the celluloses may be regarded as polymerides of several basal units, in some of which an even and in others an odd number of C_6 chains are present. The contribution of the organic chemist has at least shown that these units are exclusively composed of glucose residues, that the

linkage of these residues is in all cases identical, and that one of the units present is anhydro-triglucose.

I have merely sketched the position reached in the study of cellulose, and must hasten to the remaining polysaccharides, starch and glycogen, which display similarly the feature of being derived from glucose. It would be attempting the impossible to enumerate, in the course of this address, the points of distinction which these compounds show relative to each other and to cellulose. Both compounds give a trimethyl derivative, differing only to a small extent in specific rotation and, in each case, subsequent hydrolysis gives 2:3:6-trimethyl glucose. Now that we know where we stand regarding maltose on methylation this result is consistent, in place of being regarded as abnormal and inexplicable.

The behaviour of starch on methylation is full of instructive features, as the process can be arrested at three definite stages, and the analytical composition of these consecutive products points strongly to the idea that there are nine, or a multiple of nine, hydroxyl groups in the unpolymerised molecule. The first compound to satisfy this condition is, of course, an anhydro-trisaccharide, while the next simplest would be the corresponding anhydro-hexa-saccharide. If we refer back for a moment to the constitution of maltose and of cellobiose you will remember that we have to make a selection from two alternatives. This necessity extends likewise to the polysaccharides which yield these disaccharides, and we have already selected one alternative formula in the case of cellulose. There remains the equally arbitrary selection of the surviving option to accommodate the case of starch, which we may thus regard as composed of glucose residues joined through positions 1 and 5. At this stage the organic chemist joins forces with the biochemist, and recognises that the action of selected enzymes on starch may give either maltose or *iso*-maltose. It is thus possible to construct a formula for starch which satisfies both chemical and biochemical con-

siderations. This formula must contain an even number of C_6 residues, and the total number of hydroxyl groups must be a multiple of nine. These conditions are fulfilled by the following structure which is similar to that suggested by Ling.

XXI
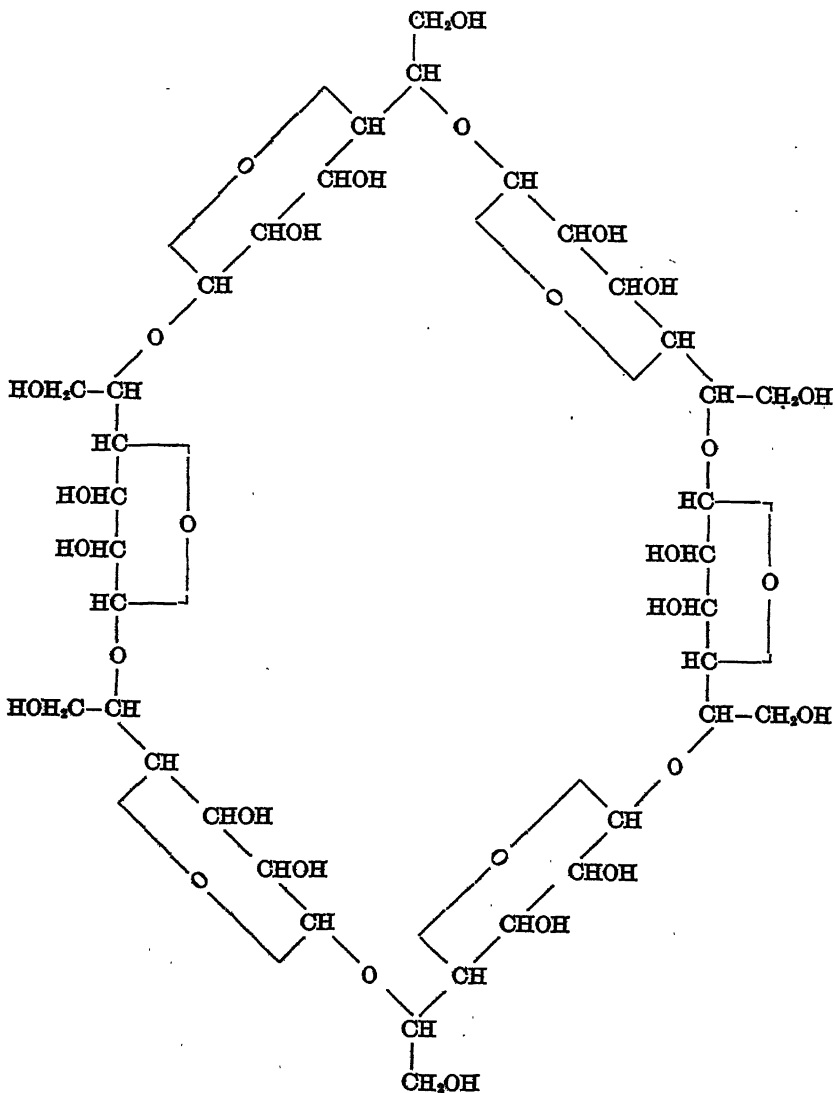
Such a formula is not founded on speculation alone. For example, it demands that out of a total of 18 hydroxyl groups twelve should be different from the remaining six, and methylation confirms this in all respects down to the actual positions occupied by these groups. Nevertheless we must be careful not to be too rigid in our views, and ought rather to remember that there is not one starch but many. We are not even justified in claiming that anhydro-hexo-glucose is the one and only molecular unit of starch, and we have no knowledge whatsoever of the different polymerides to which this unit may give rise.

The comparable researches on glycogen are concerned with minute details unsuited to a general discussion of this nature, and time permits only of brief reference to inulin, the remaining common polysaccharide. Here we are confronted with exactly the same phenomenon as we encountered in the case of sucrose, for, although inulin on hydrolysis yields laevo-rotatory fructose, this reaction provides misleading evidence as to structure. Inulin is, in fact, exclusively composed of γ-fructose residues, and treatment with acids thus initiates two consecutive reactions, the first being hydrolysis to give the unstable variety of the ketose, and the other the conversion of this sugar into the ordinary stable form. Methylation places a check on the second of these changes and, in consequence, the hydrolysis of trimethyl inulin gives a *dextro-rotatory* trimethyl γ-fructose, in which the internal oxygen ring can no longer fluctuate. This at once proves that the fructose residue in inulin is identical with that in sucrose, and we are now in a position to consider what is the simple molecule which, by polymerisation, gives the polysaccharide. The situation is a little different from that presented by cellulose and starch, as inulin differs from these polysaccharides in its failure to yield a disaccharide as the penultimate product of hydrolysis. In consequence, there is meanwhile no justification for considering any more complicated possibility than that inulin is a polymeride (or series of polymerides) of anhydro-γ-fructose, to which the following structure may be assigned:

$$
\left[\begin{array}{c}
\cdot \quad CH_2OH \\
\mid \\
\underline{}C \\
\mid \\
CH \\
\mid \\
CHOH \\
\cdot \quad O \mid \\
\underline{}CH \quad \cdot \\
\mid \\
CH_2OH
\end{array}\right]_n
$$
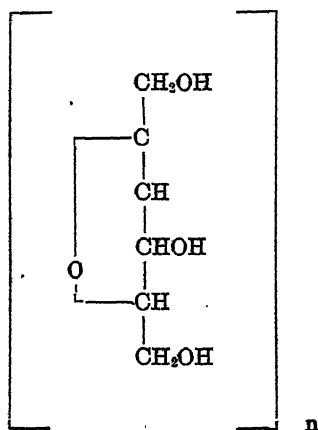
XXII

It is perhaps a simple statement to make, but it opens out vast possibilities for discussion, and raises the question as to the reasons underlying the fact that in the two great natural sources from which it is derived fructose exists in the γ-form.

At this stage, I must apply the closure to our review which, although lengthy, has included only the salient parts of the arguments upon which these conclusions are founded, and has left untouched the supplementary studies involved. These have not been haphazard or without design, as our investigations of such substances as the anilides, acetone derivatives of sugars, or of the phenomenon of optical activity or of polymerisation will show. We have been piecing together the framework of evidence on a systematic plan. The present is a time at which it is opportune to look into the future, as well as on the past, and to ask ourselves on what lines ought sugar chemistry to advance if it is to play its part in unfolding the mysteries of living chemistry. If I may venture to speculate, it would be to say that the constitutional study of carbohydrates will quickly reach its limit. We are fast approaching the point where we shall know what these substances are, without knowing why and how they react in Nature. The sugar chemistry of the future will, I trust, be concerned chiefly with syntheses and reactions conducted not

under the artificial conditions of the organic laboratory, but within the narrow range of temperature and with the modest reagents of the living cell.   It will be a chemistry controlled by severe physical restrictions, energy will be supplied by light rather than by heat, diffusion will replace filtration, polymerisation and asymmetric syntheses will be daily operations.

I began with an expression of gratitude, and, prompted by the recollection of all that has been involved in the work now laid before you, I end on the same note.   Memory conjures up the loyal collaborators, drawn from all the continents, who have come to St. Andrews to give enthusiastic help.   Nor do I forget what I owe to the discoveries which have proceeded from the research laboratories in this and other lands, for although under the special circumstances of my address I have confined myself largely to my own problems, I am deeply conscious of the fact that in scientific effort each man's work is but the making of a thread which goes to fashion the great fabric of truth.

STATIC AND DYNAMIC ISOMERISM IN PROTOTROPIC COMPOUNDS[1]

T. M. LOWRY

University Chemical Laboratory, Cambridge, England

A. THE DEFINITION OF TAUTOMERISM

1. A new definition of tautomerism

The discussion on tautomerism, at the Oxford Meeting of the British Association in August, 1926, seems likely to become of historic importance on account of the novel definition of this much-disputed term on which the discussion was based. Moreover, in view of the comments made upon it, there seems to be a very reasonable prospect that the admirable definition which was then put forward may be adopted generally, and may thus displace once for all the diverse interpretations of the term which are now in use. The definition in question describes the phenomenon of tautomerism as follows:

This term is applied to the property exhibited by certain compounds of behaving in different reactions as if they possessed two or more different constitutions; that is, as if the atoms of the same compound or group were arranged in two or more different ways, expressible by different structural formulae.

The author of this definition has not been disclosed; but since it was cited from the Oxford Dictionary, and was circulated widely in the Preliminary Program of the Oxford meeting, it may very well be referred to as the "Oxford" definition. Its principal characteristics are (i) that it abolishes finally the distinction, upon

[1] For earlier reviews of the subject see reports on "Dynamic Isomerism" (British Association Reports, Cambridge, 1904, pp. 193–224), on "Isomeric Change" (Science Progress, April and October 1909), and on "The Mechanism of Chemical Change" (Report of the Second Solvay Conference, 1925, pp. 135–178).

231

CHEMICAL REVIEWS, VOL. IV, NO. 3

which Laar insisted so strongly, between "tautomerism" and "pseudomerism", (ii) that it defines tautomerism as a purely chemical phenomenon, and thus distinguishes it sharply from the physico-chemical phenomenon of dynamic isomerism, (iii) that it abolishes the former limitation of the term to prototropic compounds and extends it to mobile structures of other types.

2. Tautomerism and pseudomerism

The term *tautomerism* was used for the first time by Laar in 1885, in a paper, "Ueber die Möglichkeit mehrerer Struktür-formeln für dieselbe chemische Verbindung" (1). This paper was based on the observation of Zincke (2) that the phenylhydrazone derived from naphthaquinone was identical with the phenylazo-derivative of α-naphthol, with which it should have been isomeric:

$$C_6H_5 \cdot NH \cdot N : C_{10}H_6 : O \qquad C_6H_5 \cdot N : N \cdot C_{10}H_6 \cdot OH$$
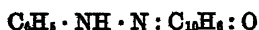
Hydrazone   Azo-compound

Since it was uncertain which of these formulae should be assigned to the identical product of the two reactions, Laar suggested that *both* formulae were equally correct, and described the supposed existence of a dual structure in the compound as "tautomerism."

The term *pseudomerism* was introduced by Laar in 1886, (3) in order to describe von Baeyer's discovery (4) that a substance, to which a perfectly definite structure could be assigned, might yield (or be prepared from) derivatives having a totally different structure. Thus, Baeyer's observations showed that, although isatin yields an O-ethylisatin by direct ethylation, an isomeric N-ethylisatin can be prepared indirectly from ethylindole by oxidation, although the pseudo-isatin from which it is derived is altogether unknown:

Isatin   Ethyl-isatin   Ethylindol   Ethyl-ψ-isatin
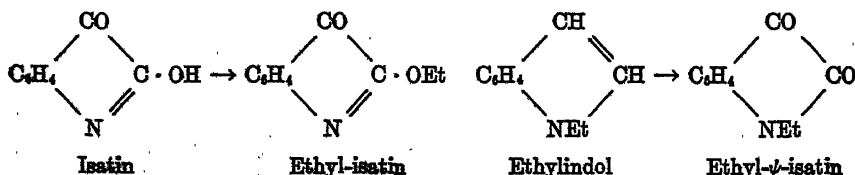
The most familiar example of pseudomerism, however, is probably that of vinyl chloride, which on hydrolysis yields acetaldehyde instead of vinyl alcohol:

$$CH_2 : CHCl \rightarrow [CH_2 : CH \cdot OH] \rightarrow CH_3 \cdot CHO$$

Vinyl chloride   Vinyl alcohol   Aldehyde
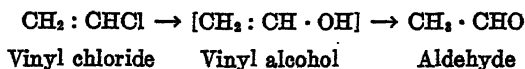
A converse case is that of camphor which, when boiled with benzoyl chloride, yields the benzoyl-derivative of an unknown enolic isomeride:

$$C_8H_{14}\Big\langle \begin{matrix} CH_2 \\ | \\ CO \end{matrix} \rightarrow C_8H_{14}\Big\langle \begin{matrix} CH \\ \| \\ C \cdot O \cdot CO \cdot C_6H_5 \end{matrix}$$
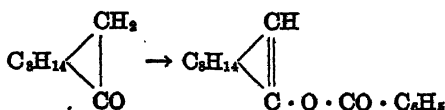
The Oxford definition, in obliterating finally the distinction which Laar made between the phenomena of pseudomerism and of tautomerism, has the advantage of being in agreement with the carefully-considered definition of Kurt Meyer (5) who uses "tautomerism" as a general term to describe the three phenomena of: (i) *desmotropy*, where the parent compound behaves as a mixture of two isomerides, (ii) *pseudomerism*, where the parent compound is of known structure, but gives rise also to a different structure in some of its derivatives, (iii) *cryptomerism*, where the nature of the underlying isomerism is still obscure. Since, after the lapse of 40 years, many of the examples cited by Laar can still only be described as "cryptomeric," there is an obvious advantage in making use of a definition which covers all these cases without attempting to classify them.

3. Tautomerism and dynamic isomerism

The necessity for a new definition of tautomerism arises from the fact that Laar embodied in his original definition a *theory* which is now universally recognized as being incorrect, at least in the cases to which he applied it, namely, that the various formulae which can be assigned to a tautomeric compound represent "not isomeric but identical substances" (6). The new definition has the advantage that there is *no* theory behind it, since it is limited

to a mere statement of the *fact* of dual reactivity, and does not attempt any explanation of that fact. Its scope and meaning are therefore quite different from those of the term "dynamic isomerism," since tautomerism is now defined as a purely chemical phenomenon, whereas "dynamic isomerism" is a physico-chemical phenomenon which may or may not be the original cause of the dual reactivity of a tautomeric compound.

The phenomenon of dynamic equilibrium between isomers, which we now describe as dynamic isomerism[2] was discovered by Butlerow in 1877, eight years before the appearance of Laar's first paper (7). Butlerow had found the first example of a reversible isomeric change in the case of the isodibutylenes.

$$CMe_3 \cdot CH = C\underset{\underset{CH_3}{\uparrow\text{------}}}{\overset{CH_3}{<}} \rightleftharpoons CMe_3 \cdot CH_2 \cdot C\underset{\underset{CH_2}{\downarrow\text{------}}}{\overset{CH_2}{<}}$$
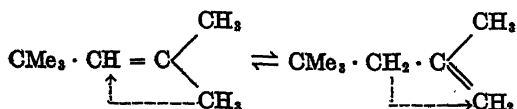
He therefore suggested that a similar condition might prevail in more labile compounds, and, in particular, that the isomeric esters of cyanic and hydrocyanic acid might be derived from real isomers, which had, however, been brought into a condition of dynamic equilibrium with one another, and could not therefore be isolated as separate entities:

$$CH_3O \cdot C \vdots N \leftarrow HO \cdot C \vdots N \rightleftharpoons H \cdot N : C : O \rightarrow CH_3 \cdot N : C : O$$

Methyl cyanate   Cyanic acid   Methyl isocyanate

$$CH_3 \cdot C \vdots N \leftarrow H \cdot C \vdots N \rightleftharpoons H \cdot N : C : \rightarrow CH_3 \cdot N : C :$$

Methyl cyanide   Hydrocyanic acid   Methyl isocyanide
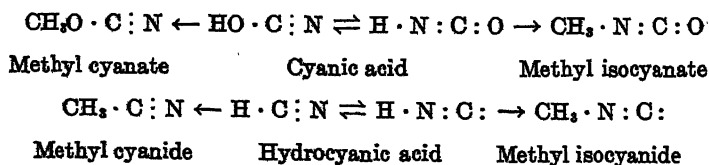
Butlerow, unfortunately, was much more concerned with the facts which he had established than with the very important theoretical conclusions which he had derived from them. He therefore published his work under the modest title "Ueber Isodibutylen," and did not even invent a term to cover the condition of *dynamic equilibrium amongst isomers* which he had discovered. Laar, on the other hand, achieved immortality because he invented an

---

[2] This term was first used in a paper on "Nitrocamphor as an Example of Dynamic Isomerism" Lowry: *J. Chem. Soc.* **75**, 211 (1899).

attractive name to describe the facts recorded by other workers, and, moreover, took the precaution of setting out the substance of his theoretical considerations in the title of his paper. As a result of this policy, the term which he introduced has come into general use, in spite of the fact that it was linked with an explanation which we know to have been incorrect, at least in the main group of cases to which he applied it.

A condition of dynamic equilibrium amongst isomers, similar to that which Butlerow discovered in 1877 in the alcohols and olefines of the C_8 series, had been postulated five years before by Kekulé (8) as a remedy for the most obvious weakness of his new formula for benzene, namely that it predicted an isomerism, which could not be realized in practice, between the 1:2 and the 1:6 diderivatives. Kekulé's dynamic hypothesis, however, had no other justification than his own reluctance to admit that carbon in aromatic compounds may perhaps be tervalent. It was therefore a mere "brain-wave," and was not supported by any trace of experimental evidence. On the other hand, since no serious chemist expects to be able to isolate isomeric compounds corresponding with all the diverse formulae which have been attributed to benzene during the last fifty years, it is clear that in this case at least we can without any hesitation apply Laar's carefully-worded definition ("Ueber die Hypothese der wechselnden Bindung" (9)), according to which the various formulae that can be assigned to a tautomeric compound represent *"not isomeric but identical substances."* Whilst therefore it would be obviously unfair to transfer from Buterlow to Kekulé the merit of discovering the phenomenon of reversible isomeric change, we may nevertheless assign to him the credit of having put forward an example of tautomerism which approaches more nearly to Laar's definition than any other case that has been investigated since the problem was first discussed.

4. Tautomerism and prototropy

Another important feature of Laar's paper is the list of *substances* in which he claimed to have detected the dual structure of a tautomeric compound. Almost all of these substances owe

their dual reactivity to a mobile hydrogen atom, the position of which is difficult to determine, either because different reactions appear to locate it in different positions in the molecule, or because it migrates very easily from one position to another even when it has been definitely located.   It is clear that neither the phenomenon of dual reactivity, nor that of reversible isomeric change, need be limited to compounds containing a mobile hydrogen atom; nevertheless, the influence of Laar's original list of cases has proved so strong that this factor is frequently included in the definition of tautomerism.   Thus, in the recent translation of Schmidt's "Text-book of Organic Chemistry," dated 1926, we read:

A substance is tautomeric when it forms two series of derivatives. These are derived from two parent structures *which differ only in the position of a hydrogen atom and of one or more double bonds.*[2]

In the same way Kurt Meyer (5) (*loc. cit*). suggests that

Substances are tautomeric, if they form two series of derivatives which are deduced from two isomeric formulae; *these formulae differ from one another in the position of a hydrogen-atom, and of one or more double bonds.*[2]

Since nobody now believes in Laar's theory of tautomerism, as he himself defined and applied it, this choice of an alternative characteristic from his schedule as the basis of a new definition is as lawful as any other; but the limitation which it imposes is out of harmony with the terms of the "Oxford" definition cited above, and, in my opinion, it should be abandoned, if only because it excludes the most characteristic of all cases of tautomerism, namely that of hydrocarbon benzene.   On the other hand, the isomeric changes which depend on the mobility of a hydrogen atom are so special and so important that it is very desirable that they should be described by some less ambiguous word, especially since it is now proposed to assign to the term "tautomerism" a meaning in which this particular factor is entirely ignored.   The

[2] The italics are mine, T. M. L.

necessity for such a term, which is admitted in both of the preceding definitions, was recognized 40 years ago by Jacobsen, who introduced the word *desmotropy* for the express purpose of describing *"a rearrangment of bonds consequent upon the displacement of a hydrogen atom."*[4] When used by Jacobsen, therefore, (10) the word has nothing whatever to do with the possibility of isolating the various isomeric hydrides. The latter alternative meaning, which was suggested by Hantzsch and Hermann in 1887 (11) has however, rendered Jacobsen's term ambiguous, and therefore useless for its original purpose. As a result, there was, until recently, no word available to describe specifically the important group of balanced isomeric changes which Jacobsen sought to differentiate. Since Jacobsen's "bond-shifting" has thus been put out of action as a description of the migration of a proton and the rearrangement of bonds which accompanies it, I have fallen back on the other and more characteristic aspect of this dual process, and have described Jacobsen's phenomenon as "proton-shifting" or *prototropy* (12). In view of the interest attaching to this type of isomeric change, and the large amount of attention that has been given to it, the remaining sections of this report are devoted to a consideration of the mechanism of prototropic change, and to the experimental study of the conditions under which it can be arrested and promoted.

B. THE MECHANISM OF PROTOTROPIC CHANGE

5. Butlerow's two types of prototropy

The readiness with which isomeric hydrogen-compounds can be converted into one another varies very greatly. Thus, in many

---

[4] The complete quotation is as follows:- "The word 'tautomerism' is based on Laar's view, which (I believe) is not shared by most chemists, that the molecules of compounds whose chemical behaviour is represented by two structural formulae differing in the point of attachment of a hydrogen atom never assume a definite constitution, but exist in a constant state of oscillatory change. The majority of chemists would explain the observations in question in this way, that the known forms of such compounds are to be represented by a definite grouping of atoms which in certain reactions passes over into an isomeric grouping by a rearrangement of bonds consequent upon the displacement of a hydrogen atom." *Ber.* 20, 1732 (1887), footnote.

cases no conditions have yet been discovered under which isomeric change can be effected, whilst in others the interconversion proceeds so readily that it has not yet been found possible to isolate any one of the components of the equilibrium-mixture as a separate entity. This contrast was clearly brought out by Butlerow, who drew a sharp distinction between the isodibutylenes, which underwent isomeric change only in presence of strong sulfuric acid, and the cyanic and hydrocyanic acids, which appeared to undergo a similar isomeric change without the deliberate addition of a catalyst. Both types of change can be represented on paper by a precisely similar scheme, depending on a direct migration of a hydrogen atom from α to a β or γ position, with a consequent rearrangement of the intermediate bonds, thus:

$$C - C = C \rightleftharpoons C = C - C \quad \text{compare} \quad O - C \equiv N \quad O = C = N$$
$$H\text{------}\uparrow \quad \uparrow\text{------}H \qquad\qquad H\text{------}\uparrow \quad \uparrow\text{------}H$$
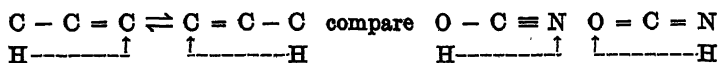
but, in spite of this superficial analogy, I have felt for many years that there is a very real difference between the two groups of cases. Further consideration has confirmed this instinctive opinion, and has led me to conclude, not only that a definite mechanism is required to bring about isomeric change in all prototropic compounds, but also that there is an important difference in the mechanism by which isomeric change takes place in Butlerow's two types of cases.

6. Mobility and acidity in prototropic compounds

The term "prototropy" is based upon the view that the wandering of a hydrogen atom in a prototropic compound is an ionic reaction, in which the mobile atom migrates as a hydrogen ion or proton, i.e., as a *positively charged nucleus* and not as an electrically-neutral atom. Since the liberation of a proton as a hydrogen ion is specially characteristic of *acids*, it is obvious that prototropic change must be closely related to the property of *acidity*. Thus, according to modern views, an acid always exists in a condition of dynamic equilibrium, in which a proton is continually being exchanged between the acid and the solvent,

e.g., $H_2O + HX \rightleftharpoons \overset{+}{H_3O} + \overset{-}{X}$. If, however, the ion X is "multipolar" (13), or is itself capable of undergoing isomeric change, so that the proton can be reattached at different points on the anion, thus giving rise to a series of isomeric hydrides, all the conditions for a reversible prototropic change are fulfilled. It is therefore probably not a mere coincidence that *prussic acid* from which, after the lapse of nearly half a century, isomeric forms have not yet been separated, is an *acid* of a more definite character than most of the compounds in which this separation has been accomplished. It is indeed obvious that a compound from which a proton can be separated by mere contact with water is in an exceptionally favourable condition for undergoing prototropic change, provided always that there are two alternative positions in which the proton can reattach itself to the anion. On the other hand, it follows from this hypothesis that, if in the future the "acidity" of the compound could be suppressed completely, e.g., by drastic drying, the mobility of the hydrogen atom would also be suspended.

Evidence of this character is cited in a later paragraph, in which it is shown that, just as acidity can only be developed in presence of a *proton-acceptor* such as water, so also prototropic change in a methylated sugar can only take place *when the medium possesses both acidic and basic properties.* Similar evidence of the close relationship between acidity and mobility in prototropic compounds is found in their sensitiveness to alkaline catalysts, since it would be difficult to find an interpretation of this effect apart from the obvious influence of a base in developing the latent acidity of the compounds upon which it acts.

7. Mechanism of isomeric change in olefines

The fact that Butlerow's isodibutylenes are entirely inert towards alkalis, but undergo isomeric change in presence of a strong acid, is *a priori* evidence that these hydrocarbons do not contain a mobile hydrogen atom of the normal type. A clear explanation of their behaviour can, however, be given if we suppose that they *acquire* a mobile hydrogen atom when the

olefine combines with the ions of sulfuric acid to form an iso-
dibutyl sulfate.[5]

$$C_8H_{16} + H_2SO_4 \rightleftharpoons C_8H_{17} \cdot O \cdot SO_2 \cdot OH$$
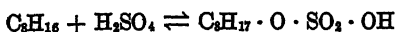
According to this view, the hydrogen attached to the α and γ
atoms of the 3-carbon system of the isomeric olefines is entirely
immobile; but this state of affairs is altered completely when the
sulfate radical is attached to the intermediate β-atom. This
strongly negative radical then acts in the usual way to promote
the separation of a positively-charged radical from the system;
it therefore makes it possible to split off a proton from *either* of
the adjacent atoms of carbon, and by a reversal of this process to
promote the migration of a hydrogen from the α to the γ position,
or conversely:
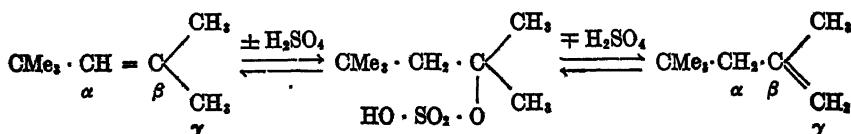
8. Absolute and conditional tautomerism

From a consideration of these examples it is clear that, whilst
the tautomerism of prussic acid is apparently *absolute* (since we
do not know any method of arresting the reversible isomeric
change, and cannot therefore hope to find any reagent which will
enable us to distinguish between the two components of the
mixture) the tautomerism of the isodibutylenes is *conditional* upon
the presence of a strong acid. It will be noticed that Laar re-
garded all cases of tautomerism as absolute, but that subsequent
workers have succeeded in finding conditions and reagents which
have enabled us to recognize in one case after another that the
tautomerism is only conditional. Thus, Knorr (14) in addition

---

[5] Butlerow had proved that the conversion of amylene into amyl hydrogen
sulfate was a balanced action; he had therefore strong experimental support for
the view that the isomeric change of the isodibutylenes was effected through the
reversible formation of a sulfate, and depended on the fact that *ter*-isodibutyl
sulfate gave *two isomeric olefines* when sulfuric acid was eliminated from it.

to separating the two components of aceto-acetic ester, was able to prove that they reacted differently towards ferric chloride, and were therefore not tautomeric towards this reagent; and Kurt Meyer (15) was able to develop a general method whereby the ketonic and enolic forms of a prototropic compound can be estimated by titrating the latter form with bromine. There can, therefore, be little doubt that further experiments will provide evidence that other cases of tautomerism are conditional upon the presence of some suitable catalyst; but, since nearly all prototropic compounds are very sensitive to the action of alkali, the reagents used in this work must be free from alkali, and in general must be neutral substances.

An interesting example of conditional tautomerism is afforded by the α and α' chlorocamphors and bromocamphors. Since these isomers are brought into equilibrium by alkalis, they must necessarily behave as tautomeric substances towards all alkaline reagents; but they are not necessarily tautomeric in acid solutions, where isomeric change is inhibited. It is therefore interesting to notice that, whilst the two chlorocamphors yield a pair of isomeric chloronitrocamphors on nitration (16)

$$C_8H_{14}\Big\langle\begin{array}{c}CHCl\\ |\\ CO\end{array} \longrightarrow C_8H_{14}\Big\langle\begin{array}{c}CCl\cdot NO_2\\ |\\ CO\end{array}$$

α and α'	$\alpha\alpha'$ and $\alpha'\alpha$
Chlorocamphors	Chloronitrocamphors
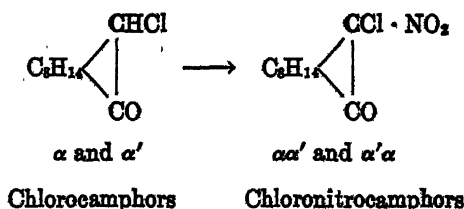
the two bromocamphors yield an identical product. Examples such as these suggest that the relationship between tautomerism and dynamic isomerism is not so close as has generally been supposed, since it is clear that dual reactivity may be developed under conditions which appear to rule out the possibility of reversible isomeric change.

C. PROMOTION AND ARREST OF PROTOTROPIC CHANGE

9. Catalysis of mutarotation by bases, acids and salts

At a date when the nature of the changes which give rise to mutarotation was still unknown, O'Sullivan and Tompson (17) recorded the fact that mutarotation takes place instantly on the addition of alkali to a solution of a reducing sugar. A similar effect was observed in the first experiments in which mutarotation was recorded as a sequel to isomeric change in a prototropic compound, since a solution of nitrocamphor, in alcohol to which sodium had been added in the proportion of 0.25 grams of metal per litre, already showed a steady rotatory power when the first reading was taken after an interval of only three minutes from the time when the solution was prepared (18). No analogous acceleration was observed, however, when a decinormal solution of hydrochloric acid in alcohol was used instead of alcohol as a solvent for nitrocamphor, although acids were known to have a marked influence in accelerating the mutarotation of the sugars. Later experiments (19) have shown that the velocity of mutarotation of an aqueous solution of glucose is a symmetrical function of the hydrogen-ion concentration, falling to a flat minimum at pH $= 5$, and rising steeply when pH is less than 2 or greater than 8. The acceleration therefore begins to be appreciable in acid solutions at concentrations above $N/100$ and in alkaline solutions at concentrations above $N/1,000,000$.

Since nitrocamphor is insoluble in water, and the values of pH in non-aqueous solutions are unknown, it is not practicable to plot a similar curve for the influence of hydrogen-ion concentration on the velocity of mutarotation of this compound. Experiments on the mutarotation of solutions of nitrocamphor in benzene (20) showed, however, that *acids* as well as bases can act as catalysts for the underlying prototropic change, although this effect was masked by the catalytic activity of the solvent when the first test was made in aqueous alcohol as described above. Experiments on the catalytic activity of acids are complicated by the fact that, even at ordinary temperatures, a further ir-

reversible isomeric change of the Beckmann type takes place, which does not occur in the presence of alkalis.

$$C_8H_{14} \diagdown \begin{array}{c} CH \cdot NO_2 \\ | \\ CO \end{array} \rightleftharpoons C_8H_{14} \diagdown \begin{array}{c} C : NO \cdot OH \\ | \\ CO \end{array} \rightarrow C_8H_{14} \diagdown \begin{array}{c} C \diagup \hspace{-2pt} \overset{NOH}{\diagdown} \\ O \\ C \diagup \hspace{-2pt} \diagdown \\ O \end{array}$$

Nitrocamphor ψ-Nitrocamphor   Camphoryloxime
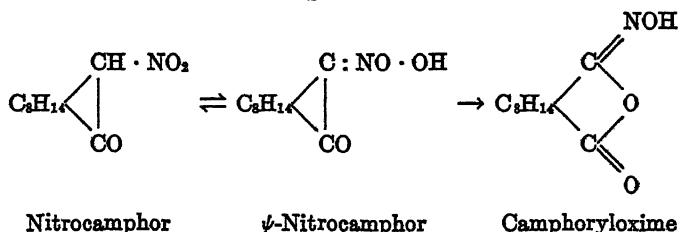
Again, when chloroform was used as a solvent, it was found that acids may retard and even arrest the mutarotation of nitro-camphor, by accelerating the oxidation of chloroform to carbonyl-chloride, which acts as an anticatalyst for the mutarotation (see below, p. 245). In spite, however, of the limitations to which experiments of this type are subject, it was shown clearly that acids as well as bases have a definite catalytic activity in promoting the mutarotation of nitrocamphor. In particular, it was found that N/10,000 trichloracetic acid produced about the same acceleration as N/100,000 piperidine.

The remarkable catalytic activity of *bases* in the mutarotation of nitrocamphor can be illustrated by the fact that the velocity of change was doubled (as contrasted with the velocity normally developed by casual impurities in the inert solvent) when the proportion of base was only one part in 10 million, or 1 decigram per ton of benzene. Piperidine in benzene was shown to be 100,000 times more active than aniline in the same solvent, and 100 times more active than sodium ethoxide in alcohol; but, since none of these effects can be attributed to hydroxyl ions, which were not present in any of the solutions, except perhaps as impurities, they cannot very well be expressed as a function of the alkalinity of the bases in water, or of the hydroxyl-ion concentration of their aqueous solutions.

Neutral *salts* also have a marked catalytic activity in the mutarotation of nitrocamphor. Thus the velocity of change of a solution in alcohol or in benzene was increased fourfold by shaking the solvent in the first case, and the solution in the second case,

with solid potassium chloride. This result follows naturally from the fact that the salts of nitrocamphor are neutral compounds, which do not readily undergo hydrolysis, and are therefore derived from an acid of sufficient strength to compete with a mineral acid for possession of a base. The sugars, on the other hand, which have no marked acid or basic properties, scarcely respond at all when experiments on mutarotation are carried out in presence of neutral salts, such as sodium or potassium chloride, even at relatively high concentrations.
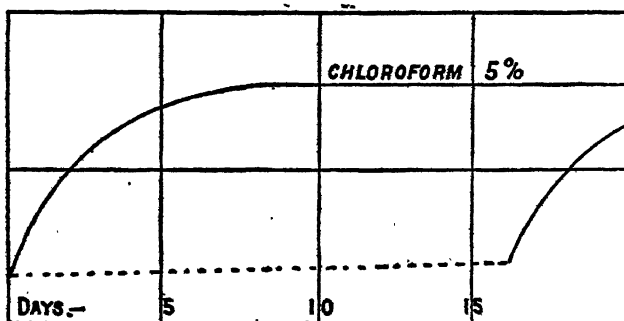
FIG. 1. CHANGE IN ROTATORY POWER OF NITROCAMPHOR IN SOLUTION IN CHLOROFORM

10. Arrest of mutarotation in inert solvents

Special interest attaches to those cases in which an isomerism, which is normally dynamic, can be rendered static by special methods of treatment, since it is very important to determine whether these facile change4 are or are not spontaneous. The first case of this kind to be recorded amongst prototropic compounds was that of nitrocamphor, where an arrest of mutarotation extending over a period of many days was observed in solutions of nitrocamphor in chloroform (21). Thus figure 1 shows that, whilst a 5 per cent solution which was transferred to a polarimeter tube of soft glass began to change immediately, and had reached a condition of equilibrium at the end of eight days, the remainder of the solution, which had been kept in a measuring-flask, had scarcely changed at all at the end of seventeen days, when it was transferred to the polarimeter tube for examination.

These abnormal solutions were probably contaminated with carbonyl chloride (formed as an oxidation product from chloroform according to the equation $2CHCl_3 + O_2 = 2COCl_2 + 2HCl$), since this compound acts as a powerful anticatalyst, especially by eliminating nitrogeneous bases in the form of carbamides. The deliberate addition of carbonyl chloride therefore made it possible at a later date to produce similar arrests of mutarotation in *benzene* and in *ether*, provided that silica vessels[6] were used to contain the solutions (22). The discovery that mutarota-
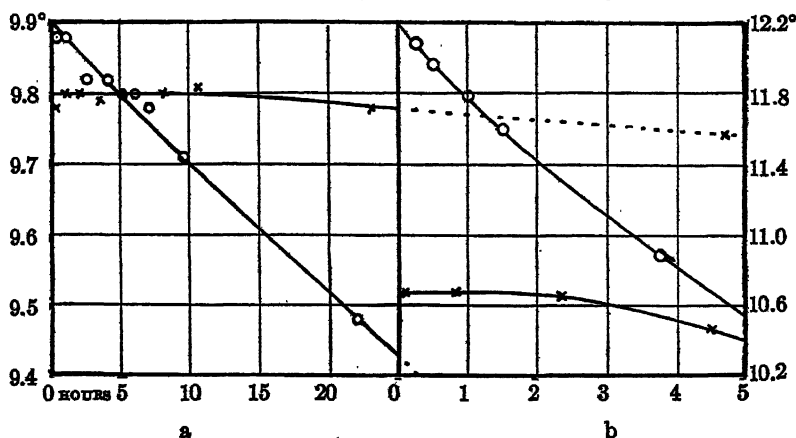
FIG. 2

a. Mutarotation of tetramethylglucose in "dry" and "wet" ethyl acetate. ×, dry. o, wet.

b. Mutarotation of tetramethylglucose and glucose in "dry" pyridine. o, glucose. ×, tetramethylglucose.

tion could be arrested almost completely in ether was specially important, since the opinion had been held previously that whilst hydrocarbon solvents were inert, oxygenated solvents in general were active promoters of mutarotation. This early opinion was also disproved when it was found that the mutarotation of tetramethylglucose in dry *ethyl acetate* (fig. 2, a) could be arrested completely during a period of at least 10 hours before a change was slowly initiated (23).

[6] Compare the later work of Kurt Meyer: *Ber.* 53, 1410 (1920); 54, 579 (1921) on the aseptic distillation of ethyl acetoacetate in vessels of alkali-free silica-glass.

Even more important was the discovery that the mutarotation of a sugar in *pyridine* can be arrested for two or three hours by moderate purification and drying (23), since the legend that pyridine was a catalytically-active solvent like water had been widely circulated and generally accepted. The complete arrest during a period of two or three hours of the mutarotation of tetramethylglucose in dry pyridine is shown in figure 2 b, which

a   b
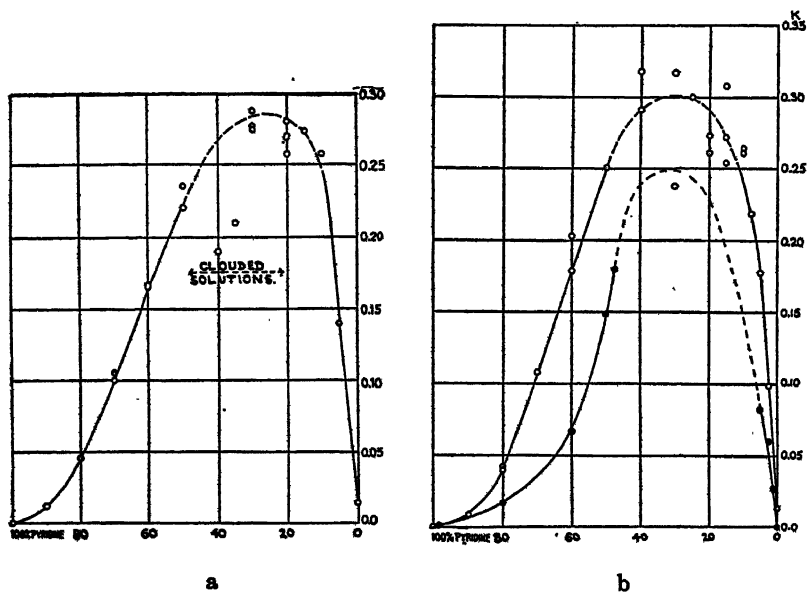
FIG. 3

a. Mutarotation of glucose in pyridine and water.
b. Mutarotation of tetramethylglucose (i) in pyridine and water ⊙, (ii) in pyridine and cresol ●.

also illustrates the failure to produce a similar arrest in the mutarotation of glucose, 'the purification of which to the point of yielding metastable solutions, in any of the limited range of solvents in which it will dissolve, is a task of much greater difficulty than in the case of the methylated sugars.

A final stage in this development of the theory of mutarotation was reached when it was shown that *cresol*, like pyridine, is not

an active catalyst when used alone, but that a mixture of cresol and pyridine, in virtue of its strongly developed amphoteric properties, is a far better catalyst than pure water (24). This is shown in figure 3 where curves are given showing the velocity of mutarotation of tetramethylglucose in mixtures of pyridine with cresol and with water, as well as for the mutarotation of glucose in aqueous pyridine.

11. Further experiments on the arrest of mutarotation

Recent experiments of the author, carried out with the collaboration of Mr. G. Owen, have been devoted to finding conditions under which these arrests of mutarotation, which were at first observed only accidentally, may be produced at will. This was done mainly with a view to determining the influence of small quantities of various catalysts *on solution which were sufficiently pure to inhibit the occurrence of isomeric change*, since under no other conditions would it be possible to guard against secondary effects produced by interaction of the added catalyst with unknown catalysts already present in the solution. The principal results of these experiments were as follows:

a. Purification of polarimeter tubes. The ordinary methods of cleaning a polarimeter tube, e.g., by washing with absolute alcohol and with water, soaking with chromic acid, and then washing repeatedly with water, leave it in a highly active condition, which gives rise to rapid mutarotation when a "clean" solution of tetra-acetyl glucose in ethyl acetate is introduced. The best method of purifying the tube is to add fresh quantities of the same "clean" solution, without any intermediate washing, when slower and slower mutarotations are produced. These remarks apply only to silica polarimeter tubes which are clean enough to give quite low velocities of mutarotation; they do not necessarily apply to glass tubes, where the alkaline silicates may perhaps act like the alkaline phosphates or borates of a buffer solution to produce a steady concentration of alkali, and may thus give rise to more uniform velocities than those which are observed when using silica tubes.

b. *Arrest of mutarotation in silica flasks.* Whilst silica polarimeter tubes can never be relied upon to give an arrest of mutarotation, much more confidence can be placed in the behaviour of *silica flasks which have been ignited and allowed to cool in a desiccator containing phosphoric oxide.* By making use of this discovery it was possible to build up a technique for studying the behaviour of solutions exhibiting arrested mutarotation. This technique depends on making up a solution *in the silica flask*, and transferring samples to a polarimeter tube (1) immediately, and (2) at the end of twelve or twenty-four hours. The solutions in the tube usually exhibit a more or less rapid mutarotation; but when, as sometimes happens, the initial readings of the two samples are identical, it is clear that the solution in the flask is not undergoing isomeric change. It is then possible to add *to the solution in the flask* a trace of a catalyst, and thus to test the effect of the addition of the catalyst *to the uncontaminated solution* by taking out another sample at the end of a further period of twelve or twenty-four hours. In view of the remarkable effects which are produced by mixing together compounds of different types which possess catalytic properties, this condition appears to be essential if trustworthy conclusions are to be drawn from the experiments.

c. *Influence of water.* A noteworthy feature of these observations was the impotence of water as a catalyst when used in presence of a large excess of ethyl acetate. Thus, in one case, in which a complete arrest of mutarotation had been revealed by the examination of two samples taken from a silica flask at an interval of fourteen hours, the addition of a drop of water to the second sample in the tube (giving a water-concentration of 0.3 per cent) produced a mutarotation with a half-change period of about seventy days, as compared with seventy-five days for the original dry sample. The inability of water to promote mutarotation, when added to a hygroscopic medium which possesses no inherent catalytic properties, was in agreement both with our own theoretical views and with recent observations on the mutarotation of tetramethyl-glucose in aqueous acetone (25) which showed that mutarotation was almost completely inhibited in solutions containing less than 5 per cent of water. It is indeed obvious

that, in experiments of this kind, it is far more important to make use of clean apparatus than of perfectly dry solutions, and that the object to be attained is not so much "Bakerian dryness" as "Bakerian cleanness."

d. *Influence of dilute acids and alkalis.* In complete contrast to the impotence of pure water was the effect produced by the addition of a drop of decinormal hydrochloric acid to a solution, which was showing an almost complete arrest of mutarotation *in the polarimeter tube.* In this case, a half-change period, which had gradually increased to one hundred and twelve days, fell to six hours on the addition of 0.3 per cent of water containing 0.3 per cent of HCl. A final acid-concentration of only 0.001 per cent had therefore increased the velocity of change about 500-fold, although pure water in a duplicate experiment appeared to retard rather than to accelerate the action. In a precisely similar way the addition of a single drop of normal sodium hydroxide to a polarimeter-tube, containing a solution which was giving a half-change period of about one hundred days, again showed a 500-fold acceleration, the half-change period falling to about five hours. It is noteworthy that this addition of alkali "fouled" the polarimeter tube to such an extent that, when a further sample of the clean solution from the flask (which was still showing a half-change period of about sixty-four days) was poured into the tube, it exhibited an almost identical velocity of mutarotation, with a half-change period of about six hours.

D. CONCLUSIONS

12. *"Tautomeric" changes not spontaneous*

When Laar put forward his theory of "tautomerism," he supposed that the migration of a hydrogen atom was a form of perpetual motion, like the movement of a planet, and that the formulae which assigned certain fixed positions to the hydrogen atom might be compared with a record of the phases of the moon. This view became untenable, and the theory of tautomerism (but not the name) was generally abandoned, when in one case after another isomeric hydrides were separated as definite en-

tities from the equilibrium-mixtures which are formed in the liquid state or in solution. Nevertheless, the assumption, first made by Butlerow, that prototropic change is spontaneous has persisted even to the present time, since the view has been widely expressed that "tautomerism" is an intramolecular change, in which the solvent does not intervene (26).

It cannot be stated too clearly, however, that this primitive view, although quite plausible when first advanced by Butlerow in 1877, has been experimentally untenable since 1899, when an arrest of prototropic change was first recorded. The fact that these changes are not spontaneous, but depend on a definite mechanism, in which the molecules or ions of the medium (as well as the molecules or ions of the prototropic compound) play an essential part, has been confirmed by many subsequent examples of the arrest of isomeric change under "aseptic" conditions. It may indeed now be accepted as a general proposition that, since prototropic change can be arrested by careful purification, just as other chemical changes can be stopped by intensive drying, a definite mechanism is needed in each case, and that unless this is provided the isomerism will remain static instead of becoming dynamic.

13. The mechanism of prototropic change

The simplest basis for such a mechanism is to suppose that the migration of a hydrogen atom in a prototropic compound is an ionic process, and therefore depends on making use of a medium which possesses a dielectric constant of sufficient magnitude to enable it to act as an *ionising solvent*. According to this view, prototropic change should be possible in any solvent in which an anhydrous acid develops an appreciable electrolytic conductivity. This supposition, although it accounts for the special readiness with which mutarotation takes place in aqueous solutions, was disproved when it was found that mutarotation could be arrested not only in chloroform and in benzene, but also in ether, ethyl acetate, acetone and pyridine. A more plausible explanation of the efficiency of water in promoting the mutarotation of the sugars

was then found in its behaviour as an *amphoteric solvent*. This explanation had the advantage of assigning a normal rôle to acid and basic catalysts, since their efficiency could then be attributed to their characteristic behaviour as proton-donors and as proton-acceptors respectively. The conclusion that prototropic change depends on the transfer of a proton *to the medium*, and the recovery of a proton *from the medium*, was, however, finally vindicated, when it was shown that cresol and pyridine were impotent to promote the isomeric change of tetramethyl glucose, but that a mixture of the weak acid and base possessed a remarkably high catalytic activity.

14. Outstanding anomalies

The mechanism suggested above accounts for almost the whole range of observed facts, and there is therefore all the more interest in recording two anomalies which still remain unexplained. The first of these has reference to the fact that solutions of tetra-acetylglucose in ethyl acetate, in which mutarotation had been arrested by working under clean conditions, gave inflected mutarotation curves when a change of rotatory power was finally initiated by the addition of a drop of dilute acid or alkali. Curves of this type are often observed in solutions which undergo a progressive change as the result of a gradual absorption of impurities from the containing vessels, or the like; but it is altogether exceptional for this effect to occur when the solvent is already a complete catalyst. It is possible that, under the rather exceptional conditions of these experiments, successive stages of the process were disclosed, which are usually concealed beneath the apparently unimolecular character of the mutarotation curves.

A second anomaly is found in the fact, first revealed by Hudson, but fully confirmed by the subsequent work of Kuhn and of Euler, that mutarotation at the isoelectric point proceeds much faster than can be accounted for by allowing for the concentrations of hydrogen and hydroxyl ions in the solution. The ratio (about 10 to 1) between the observed and calculated velocities is far too great to be accounted for by any minor error, and could not

possibly be corrected by any minor alteration such as the substitution of "activity" for "concentration." Allowance can be made for it, however, by introducing a term representing the concentration of neutral water molecules in the solution. In view of the fact that ether, which resembles neutral water in its oxonium reactions, etc., is not a catalytic solvent, it is difficult to see what rôle these neutral molecules can play in promoting mutarotation except by giving rise to hydrogen and hydroxyl ions. A complete solution of this problem may, however, be looked for in the development, on lines that have already been indicated (27) of an electrolytic theory of catalysis, based upon Armstrong's theory of chemical change.

REFERENCES

(1) LAAR: *Ber.* **18**, 649 (1885).

(2) ZINCKE: *Ber.* **17**, 3030 (1884).

(3) LAAR: *Ber.* **19**, 730 (1886).

(4) VON BAEYER: *Ber.* **16**, 2188 (1883).

(5) MEYER, KURT: *Ann.* **398**, 63 (1913).

(6) LAAR: *Ber.* **19**, 730 (1886).

(7) BUTLEROW: *Ann.* **189**, 44-83 (1877).

(8) KEKULÉ: *Ann.* **162**, 86 (1872).

(9) LAAR: *Ber.* **19**, 730 (1886).

(10) MEYER AND JACOBSEN: Organische Chemie, 2 Aufl. 1, 2, 886, 910, 927.

(11) HANTZSCH AND HERMANN: *Ber.* **20**, 1802 (1887).

(12) LOWRY: *J. Chem. Soc.* **123**, 828 (1923); Second Report of the *Institut International de Chimie Solvay*, (1925), "Le Mechanism de la Transformation Chimique," p. 150.

(13) LOWRY: *J. Chem. Soc.* **123**, 828 (1923); *Bull. Soc. Chim.* [iv], **35**, 914 (1924).

(14) KNORR: *Ber.* **44**, 1138 (1911).

(15) MEYER, KURT: *Ber.* **44**, 2718 (1911).

(16) LOWRY, STEELE, AND BURGESS: *J. Chem. Soc.* **121**, 633 (1922).

(17) O'SULLIVAN AND TOMPSON: *J. Chem. Soc.* **57**, 920 (1890).

(18) LOWRY: *J. Chem. Soc.* **75**, 221 (1899).

(19) HUDSON: *Z. physik. Chem.* **44**, 487 (1903); *J. Am. Chem. Soc.* **29**, 1572 (1907); **39**, 320 (1917).

    KUHN: *Z. physik. Chem.* **113**, 389 (1924); **114**, 44 (1925).

    EULER: *Z. anorg. allgem. Chem.* **146**, 45 (1926); **152**, 113 (1926).

(20) LOWRY AND MAGSON: *J. Chem. Soc.* **93**, 117 (1908).

(21) LOWRY: *J. Chem. Soc.* **75**, 220 (1899).

(22) LOWRY AND MAGSON: *J. Chem. Soc.* **93**, 119 (1908).

(23) LOWRY AND RICHARDS: *J. Chem. Soc.* **127**, 1398 (1925).

(24) LOWRY AND FAULKNER: *J. Chem. Soc.* **127**, 2883 (1925).
(25) JONES AND LOWRY: *J. Chem. Soc.* **129**, 720, (1926).
(26) BAKER, INGELD AND THORPE: *J. Chem. Soc.* **125**, 268 (1924).
(27) LOWRY: *J. Chem. Soc.* **127**, 1383 footnote (1925).

THERMAL EQUILIBRIUM OF ELECTRONS IN METALS: CONTACT POTENTIALS AND THERMOELEC- TRIC FORCE

WORTH H. RODEBUSH

Laboratory of Physical Chemistry, University of Illinois

The electrical behavior of metals is fundamental in any theory of electricity and until recently it has been a subject of much mystery and dispute. Thus the location of the e.m.f. in the galvanic cell has been a matter of controversy since the time of Volta and Faraday. Likewise in regard to the so-called "contact potentials," there were not only differences of opinion as to the interpretation but not even general agreement as to the experimental evidence. Recent experimental work of Millikan (1) and others (2) on the photoelectric effect and the brilliant theoretical papers of Schottky (3) and others (4) on the subject of electron emission have thrown much light on the problem. While we still know less about the structure of metals than about any other solid form of matter we can at least settle some of the mooted questions of the past and formulate the problem in the case of others.

THE PHOTOELECTRIC EFFECT AND THE THERMIONIC WORK FUNCTION

The electrical behavior of metals becomes more intelligible if we first get clearly in mind the experimental facts of the photo-electric effect. When light is allowed to impinge on the clean surface of a metal in a vacuum there is a limiting frequency ν_o below which electrons are not emitted by the metal even in the presence of an accelerating field. Light of higher frequency ν ejects electrons with a kinetic energy $\frac{1}{2}mv^2 = h\nu - h\nu_o$. In accord with the Einstein law the work of removing an electron from the metal is seen to be $h\nu_o$. ν_o the limiting frequency varies widely for different metals and is evidently one of the most fundamental characteristic properties of a metal.

Electrons may be emitted by a metal without the action of light. If the metal be heated to temperatures sufficiently high, electrons acquire kinetic energy in excess of the quantity $h\nu_0$ and escape. This process is analogous to the evaporation of atoms. It is beyond the scope of this paper to discuss the phenomena of the vaporization of electrons especially as the subject has been treated at great length by Richardson (5), Schottky (3), Dushman (6) and others (7). By making certain assumptions the vapor pressure of electrons from a hot metal may be treated by the laws of thermodynamics and the heat of vaporization may be calculated by familiar formulae. The value obtained in this way appears to be in good agreement with $h\nu_0$ as determined by photoelectric measurements (8). This heat of vaporization is usually expressed in volts and designated as ϕ, the "thermionic work function."

Schottky has analyzed the work of removal of an electron from a metal as being due to various effects such as the removal of the electron from the "structure" of the metal, the overcoming of fields due to a polarization layer at the surface and electrical image attraction after the electron is through the surface. While it is not possible to measure these effects separately we may suspect that variation in ϕ for different metals is largely due to the variation in the first effect.

According to the current theories of atomic structure the peculiar properties of a metal are due largely to the so-called free or conducting electrons which are presumably identical with the valence electrons. These electrons move in orbits which lie for the most part on the outside of the atom or if they do not move in orbits they occupy the outermost energy levels in the periphery of the atom. When the atom is in the metallic lattice these electrons pass from one atom to another without appreciable energy change. When the atom is isolated, as in the vapor state, the work of removing an electron to an infinite (practically a very short) distance from an atom can be measured by the determining the ionizing potential. Obviously in the theory of electricity in metals, the ionizing potential (22) of the atom is a still more fundamental quantity than the thermionic work function.

Presumably if we knew the configuration of the atom we could calculate the ionizing potential from Coulomb's law and simple mechanics. At any rate we may assume that the ionization potential varies directly as the effective nuclear charge on the atom and inversely as some power of the effective radius. Our knowledge of atomic structure confirms this generalization.

Of course when the atoms are packed into a lattice this work of removal of the electrons will be considerably changed in magnitude. Our knowledge of the mechanics of the lattice structure is not sufficient to make any predictions here although Born (9) has made progress along this line. It seems clear however, that with a better knowledge of atomic structure and the dynamics of crystal lattices we should be able to calculate the thermionic work function for any metal. In the accompanying table are given in volts the ionizing potential, photoelectric "work function" $h\nu_o$, and the standard electrode potential against hydrogen for the three metals for which satisfactory figures for all three quantities appear to be obtainable.

	IONIZING POTENTIAL	$h\nu_o$	E_o
Na	5.13	1.7	2.71
Li	5.36	2.36	2.96
Hg	10.2	4.52	−0.80

The correspondence of the values given is apparent. It is evident that the "affinity" of the atom for the "free" electron is fundamental in determining the electrical properties of the metal. A relative measure of this property is given by various quantities such as the ionizing potential of the vapor, the limiting frequency of the photoelectric effect or the thermionic work function. If the metals be arranged in a series according to the values of any of these quantities an order will be obtained which will be approximately that of the familiar electromotive series of the electrochemist.

Now if the conducting electrons lie on different energy levels in different metals it is clear that when two metals are brought in contact, the electrons will tend to pass from one metal to

the other the motion being in every case from the metal with the smaller value of $h\nu_o$ to the one with the larger value, and this process will continue until equilibrium is established. At low temperatures where the electrons do not have appreciable kinetic energy (10) the result is easily stated. It is a fundamental law of electrostatics that a system of electrical charges tends to take on a configuration of minimum potential energy. Electrons will pass from one metal to the other until the negative potential acquired by the metal having the higher value of ν_o balances the difference in the energy levels of the electrons in the two metals.

One very peculiar point needs to be noticed here. The transfer of electrons is entirely on the surface of the two metals. No transfer takes place within the body of the metal and it is impossible to produce or maintain a volume charge within a metal. For electrons free to move but without kinetic energy it can be demonstrated from Coulomb's law that the net charge on any volume in the body of a conductor will not differ appreciably from zero if the volume is taken large enough to contain a considerable number of atoms. Even if the electrons possess kinetic energy it has been shown by Lorentz (11) that the concentration of electrons in the interior of a metal remains constant and equal to the number of positive charges in the same area. It is only within a distance from the surface of the metal comparable to the atomic diameters that appreciable changes in electron concentration can be produced even by the application of the highest potentials available.

THE THERMAL EQUILIBRIUM OF ELECTRONS BETWEEN METALS

So long as the electrons do not possess an appreciable kinetic energy the ordinary laws of electrostatics would suffice to calculate the conditions for equilibrium between metals, provided of course that we were sufficiently well acquainted with the structure of the metals. At higher temperatures where the electrons begin to share in the kinetic energy of the metal the laws of electrostatics are no longer sufficient to determine equilibrium but we must make use of thermodynamics. Equilibrium will

be established between the electrons of two metals a and b when the escaping tendency of the electrons is the same from each metal (12). The escaping tendency of any constituent from a phase is measured by the partial molal free energy of the constituent in that phase. The partial molal free energy of the electrons in a metal is the ratio $\dfrac{\partial F}{\partial n}$ where ∂F is the increase in the total free energy of the metal on the addition of ∂n equivalents of electrons. Other factors as temperature are constant. Equilibrium between the two metals a and b then may be attained either by direct contact or through the vaporization and condensation of electrons and the condition for equilibrium is

$$\frac{\partial F_a}{\partial n} \qquad \frac{\partial F_b}{\partial n} \tag{1}$$

So far, the condition for equilibrium appears the same as for the distribution of a solute between two immiscible solvents but certain important differences need to be considered. The experimental measurement of the partial molal free energy of a constituent of a solution involves the change of concentration of that constituent and we do not know of any way to change the electron concentration inside a metal, as was pointed out above. This does not invalidate the thermodynamic formula however. A more serious complication arises because of the charge carried by the electron. When we transfer electrons from one metal to another we leave a positive charge behind and carry the negative charge against the electrostatic attraction and do work so long as we increase the separation. This action at a distance makes the case quite different from the separation of a neutral molecule of a solute from a solution when the forces cease to act as soon as the molecule is separated from the surface of the solution by a distance of the order of the molecular diameter. In order to take account of the charge of the electron it is necessary to follow the method of Gibbs (13) and Schottky (3) and separate the partial molal free energy into two terms

$$\frac{\partial F}{\partial n} = \bar{F} - NeV \tag{2}$$

$\overline{F}$ is what Gibbs has called the intrinsic free energy corresponding to the free energy of a neutral molecule and V is the electrostatic potential. N is Avogadro's number, e the charge on the electron, while the minus sign takes care of the negative charge. This separation of the free energy into terms involving forces that act at molecular distances and forces acting at greater distances seems from one point of view quite arbitrary and meaningless, but it is justified by several considerations. V, the electrostatic potential, is the same for positive charges as for negative and so long as the distribution of charges does not change it is independent of temperature. $\overline{F}$ on the other hand depends upon the potential and kinetic energies of the electron in the lattice, is a function of the temperature, and the value for the electron is radically different from the value for a positive ion. Finally in some cases at least the value of V can be measured experimentally. The equation (1) for equilibrium of electrons between two metals at the same temperature then becomes

$$\overline{F}_a - NeV_a = \overline{F}_b \perp NeV_b \tag{3}$$

The metal with the larger value of the thermionic work function may be expected to have the lesser value of $\overline{F}$, lesser being used in the algebraic sense.

THE VOLTA DIFFERENCE OF POTENTIAL

The fact that electrons will pass from one metal to another on contact was observed by Volta near the beginning of the last century. The phenomenon was studied in detail by Lord Kelvin (14). In his method the two metals were made the plates of a condenser and an e.m.f. applied between of such magnitude and direction that no charge appears upon the condenser. By transforming equations (3) we have

$$V_b - V_a = \frac{\overline{F}_b - \overline{F}_a}{Ne} \tag{4}$$

It is obvious that the e.m.f. applied in the Kelvin experiment must be equal to $V_b - V_a$ in equation (4) and hence the latter term is the Volta difference of potential. Furthermore any tendency

to transfer electrons is from the surface of one metal to the surface of the other; hence the values of $\bar{F}_a$, $\bar{F}_b$ involved are the values for the metallic surfaces.

In the discussion of the photoelectric effect we assumed the existence of clean metallic surfaces but it is not certain that it is possible to obtain such a surface experimentally. Millikan in his work on the photoelectric effect of the alkali metals, shaved the surface of the metal in a high vacuum, but in the highest vacuum obtainable the surface of the metal would in all probability be quickly covered with a layer of gas molecules. Probably the cleanest surface that has been obtained is that of flowing mercury (15). It has already been emphasized that metals acquire charges only upon the surface and that in equation (4) the values of $\bar{F}_a$ and $\bar{F}_b$ are for the electrons in the surfaces. Impurities adsorbed on the surface of a metal affect profoundly the values obtained experimentally for ϕ_1 ν_0 and the Volta effect by the Kelvin method. Moreover, even if clean surfaces are obtained it is not likely that the value of $\bar{F}$ for electrons in the surface is the same as for the interior of the metal. Hence the Volta effect as measured by the Kelvin method is a superficial property of a metal and often without significance. It should be emphasized that the Volta effect is a difference of potential and not an e.m.f. The e.m.f. is applied by the experimenter. The case is analogous to the definition of osmotic pressure, where the pressure is one that is imagined to be applied rather than one that actually exists.

<center>CONTACT POTENTIALS</center>

If two metals are placed in contact without any e.m.f. applied between them, the electrons flow from one to the other in the direction of decreasing free energy until equilibrium is established according to equation (4), the difference in electrostatic potential being equal and opposite to the difference in intrinsic free energies. Between the surfaces of the metals which are not in contact there will exist a field due to the potential difference $V_b -$ V_a where $V_b - V_a$ is numerically equal to the Volta effect. This field it should be noted is external to the metals.

The difference of potential between the interior of one metal and the interior of another is given by equation (4) provided $\bar{F}_a$, and $\bar{F}_b$ are the values for the interior of the metals. Schottky has called this the "galvanic contact potential" and this is the true contact difference of potential encountered by a current flowing across the junction. It was the opinion of Volta that the e.m.f. of a galvanic cell was due to this difference of potential and resided at the junction and this opinion has persisted to the present (16). It must be obvious on consideration however that this difference of potential is exactly cancelled by the difference
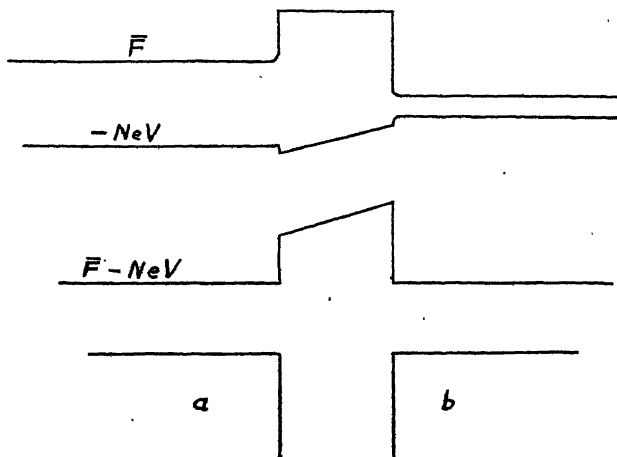
FIG. 1

in intrinsic free energies. There is no method known for measuring this true contact potential although we shall probably be able to calculate its value when we know more about the structure of metals.

The persistence of the idea that the e.m.f. of the galvanic cell is located at the metallic junction is probably due to the fact that the values of the Volta effect for pairs of metals seem to correspond in some cases at least to the differences in the standard electrode potentials of these metals. While there can be no direct relation between the "superficial" Volta effect and the

electrode potentials it would not be surprising if some parallelism were found since both depend upon the fundamental electrical properties of the metal.

In figure 1 are represented diagrammatically the course of the values of $\bar{F}$, $-NeV$, and the "total" free energy, $\bar{F} - NeV$, for the electrons of two metals a and b which are in thermal and electrical equilibrium.   The value of F_b is less than F_a and the values of $\bar{F}_b$ and $\bar{F}_a$ are assumed to rise near the surface of the metals. The value of the total free energy is the same throughout the conducting regions of both metals as is required for equilibrium of "free" electrons.   Hence there can be no e.m.f. at the junction of two metals so long as the metals are at uniform temperature throughout.

SINGLE ELECTRODE POTENTIALS

If a piece of metal is brought into contact and equilibrium with a solution containing its ions the chemical reaction may be written
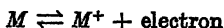
$$M \rightleftharpoons M^+ + \text{electron}$$

For equilibrium we have

$$F_M = \bar{F}_{M^+} + NeV_{sol} + \bar{F}_e - NeV_M \tag{5}$$

F_M is the molal free energy of the metal and $\bar{F}_{M^+}$ and V_{sol} are defined for the ions in solution in the same way that $\bar{F}_e$ and V were defined for the electrons in the metal.   Transforming this equation

$$V_{sol} - V_M = \frac{F_M - \bar{F}_{M^+} - \bar{F}_e}{Ne} \tag{6}$$

It has been commonly assumed that the free energy term on the right of equation (6) is of profound significance and many attempts have been made to measure $V_{sol} - V_M$.   On careful consideration the importance of this measurement is not so clear, supposing it could be made with accuracy.   The process actually is the removal of an atom from the lattice, the transfer of an electron from the ion to the surface of the metal and the solvation of the ion resulting.   This does not appear to be the exact equivalent

of one half the reaction in a cell where the electrons are transferred from one metal to the other instead of accumulating on the surface.

Granting that this somewhat complicated process is interesting it is certain that the accurate measurement of $V_{sol} - V_M$ presents enormous difficulties. We cannot measure the Volta difference of potential between two metals with certainty and the measurement must become still more difficult between a metal and a solution. In the former case only neutral molecules are adsorbed on the surface but in the latter ions may be adsorbed. Thus the hydrogen ion concentration may have a profound effect on the value of $V_{sol} - V_M$. At any rate we have a quantity here which will probably be calculated eventually more accurately than it can be measured.

THE TEMPERATURE COEFFICIENT OF CONTACT POTENTIAL

If we differentiate equation (4) with respect to temperature we obtain by well known relations of thermodynamics

$$\frac{\partial(V_b - V_a)}{\partial T} = \frac{1}{Ne}\left(\frac{\partial \overline{F}_b}{\partial T} - \frac{\partial \overline{F}_a}{\partial T}\right) = \frac{1}{Ne}(\overline{S}_a - \overline{S}_b) \qquad (7)$$

Here $\dfrac{\partial(V_b - V_a)}{\partial T}$ the temperature coefficient of the difference of electrostatic potential between the interior of two metals is seen to depend upon the respective partial molal entropies of the electrons in the interior of the metals.

THE THREE WAYS OF DEFINING THE HEAT CAPACITY OF THE ELECTRON

If we differentiate equation (7) a second time with respect to the temperature and introduce the relation $\dfrac{\partial S}{\partial T} = \dfrac{C_p}{T}$ we obtain

$$\frac{\partial^2 \overline{F}_b}{\partial T^2} - \frac{\partial^2 \overline{F}_a}{\partial T^2} = \frac{1}{NeT}\left(\frac{\partial C_{pa}}{\partial n} - \frac{\partial C_{pb}}{\partial n}\right) \qquad (8)$$

Here $\dfrac{C_p}{\partial n}$ is the partial molal heat capacity of electrons in the

metal, a quantity which cannot be measured experimentally because we cannot change the concentration of electrons in a metal except on the surface and there the change is too small to affect the heat capacity measurably.

Certain of the metals show values for the atomic heat capacity considerably above that predicted by the Debye equation. This is especially true of the metals with small ionizing potentials, such as potassium at ordinary temperatures and of many of the metals at higher temperatures. G. N. Lewis (17) has attributed this abnormally high heat capacity to the presence of electrons in an unusually free condition so that they share in the equipartition of energy. We may designate this excess heat capacity over the normal value as Ce, the "apparent heat capacity of the electrons in the metal."

Finally there is the Thomson effect. If a current flows along a wire in a thermal gradient, in addition to the joule heating there is a small heat effect that is proportional to the quantity of electricity which flows, and to the temperature change. If the current is reversed the heat effect is reversed. This is called the Thomson effect σ and may be defined as the heat absorbed per equivalent per degree rise in temperature when the electron current flows from a lower to higher temperature.[1]

Before we can discuss the possible relations of these heat capacities defined in different ways it is desirable to discuss the application of thermodynamics to the thermocouple.

THE THERMOCOUPLE AS A CARNOT CYCLE

If we neglect the irreversible flow of heat which always takes place in a thermal gradient because of the conductivity of matter, we may treat a thermocouple whose junctions are at different temperatures as a Carnot cycle. Whether we are justified in this procedure, in other words, whether the laws of thermodynamics apply strictly to thermocouple or not, is a question which has never been settled although most writers on the subject have inclined to the affirmative. The matter cannot be settled by

---

[1] The Thomson effect has usually been defined for a positive current.

an experimental check because the heat quantities involved are too small to measure with accuracy. It may be pointed out that no Carnot cycle can be carried out without some irreversible flow of heat for we have no perfect insulators. In the ordinary Carnot cycle, however, the thermal gradient may be assumed to be external to the mechanism while in the thermocouple the electrons move through the thermal gradient. For purposes of discussion in the remainder of this paper we shall assume that the thermocouple may be treated as a Carnot cycle.

Let us consider a circuit of two metals a and b in a temperature gradient with one junction at the temperature T and the other junction at the temperature $T + dT$. Assuming the electrons to flow from a to b at the warmer junction then the net electromotive force of the circuit dE in the direction of the electron current is given by the First Law as

$$Ne \, dE = q_2 - q_1 + (\sigma_a - \sigma_b)dT \qquad (9)$$

Here q_2 and q_1 are the heats absorbed at the warmer and colder junctions respectively when one equivalent of electrons flows from a to b and σ_a, σ_b are the Thomson effects. By the Second Law:

$$Ne \, dE = \frac{q_2 dT}{T} \qquad (10)$$

T being the temperature of the warmer junction. The Thomson effects do not appear in the above equation since they would be of second order. Differentiating (10) with respect to T and combining with (9) we have

$$Ne \, \frac{\partial^2 E}{\partial T^2} = \frac{1}{T} (\sigma_b - \sigma_a) \qquad (11)$$

We shall use this equation in the next section.

It is important to note that in equation (10) there is no information as to the numerical value of an e.m.f., E, which many writers have assumed to exist at the junction. Here q is the reversible heat and it is well known that in the analogous case of a chemical reaction the reversible heat of a reaction bears no relation to the free energy. Nevertheless the list of writers who

have sought to set q equal to E contains some of the distinguished names of science.   Furthermore it should be noted that the existence of a reversible Thomson effect does not imply the existence of an e.m.f. along the gradient.   The laws of thermodynamics give only the total e.m.f. of the circuit without giving us any specific information as to the way this e.m.f. is distributed.

One inference however, may be drawn as to the relation between the Peltier heat, q, and the e.m.f. at the junction.   If we were to tabulate the reversible heats for a number of chemical reactions, while there would be no correlation we should expect that the reversible heats would be on the average of the same order of magnitude as the free energies.   It has been a source of con-
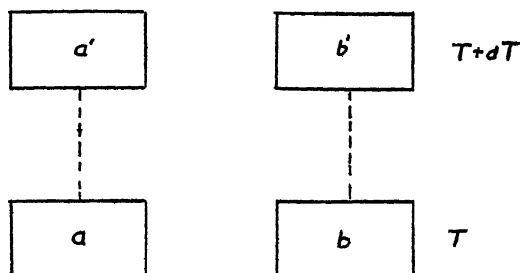
FIG. 2

cern to many writers that the values of q, as calculated from measurements of thermoelectric force, were so small compared with the Volta difference of potential.   This ceases to be a matter of concern when we recognize that the Volta effect has no relation to the e.m.f. at the junction of two metals.

A POSSIBLE CORRELATION OF THE THERMOELECTRIC POWER WITH THE "TRUE" CONTACT DIFFERENCE OF POTENTIAL

Let us consider two blocks each of metals a and b, designated as a, a' and b, b'; a and b are at the temperature T and a' and b' at the temperature $T + dT$.   Suppose each of the four pieces of metal to be electrically neutral and at the same electrostatic potential V.   The values of the intrinsic free energies will be

$\overline{F}_a, \overline{F}_b$ and $\overline{F}_a' = \overline{F}_a + \dfrac{\partial \overline{F}_a}{\partial T} dT, \overline{F}_b' = \overline{F}_b + \dfrac{\partial \overline{F}_b}{\partial T} dT.$ Now suppose a, a' and b, b' to be joined by thin wires of the respective metals a and b. These wires will be in the temperature gradient. Let us assume that no transfer of electrons takes place along the wires. The plausibility of this assumption we will discuss later. If there be no tendency for electrons to move along the temperature gradient then we may suppose that no work will be required to move electrons along the wires. Let us imagine a transfer of N electrons to take place around the circuit a a' b' b. If the electrostatic potential is the same throughout the only work involved will be in the transfer of the electrons across the gaps a' b' and ba. The net work is seen to be

$$Ne\, dE = \left(\frac{\partial \overline{F}_a}{\partial T} - \frac{\partial \overline{F}_b}{\partial T} \right) dT \tag{12}$$

If we differentiate this equation with respect to T and combine with (8) and (11) we have

$$\frac{\partial C_{pb}}{\partial n} - \frac{\partial C_{pa}}{\partial n} = \sigma_b - \sigma_a \tag{13}$$

From this equation we may infer that[2]

$$\frac{\partial C_{pa}}{\partial n} = \sigma \tag{14}$$

Latimer (19) in an interesting paper has attempted to correlate the values of Ce the apparent heat capacity of the electrons with σ but the experimental data he considers do not give a very satisfactory correlation. His conclusion would also imply that Ce is equal to $\dfrac{\partial C_p}{\partial n}$ if equation (14) is true. If we consider the case of a solution as an analogy the partial molal heat and the apparent molal heat only have the same values in general in case the solution is thermodynamically "perfect."

[2] Schottky (3) has demonstrated equation (14) by a different process of reasoning but making similar assumptions to those made here. It may be noted that the author (18) had previously called attention to the possibility that $\dfrac{\delta C_p}{\delta n}$ and σ were identical.

None of the above conclusions are true of course unless our assumption that no variation of electrostatic potential exists along a wire in a temperature gradient is correct.[3] This assumption appears very improbable. We can prove nothing from thermodynamics since the relations of the free energies are without significance unless the system is at constant temperature and recent work on the Soret effect (21) throws no light on the laws of equilibrium in thermal gradients. No doubt the values of $\overline{F}_a$ and $\overline{F}_a{}'$ in the interior of the metal will be the same after a and a' are connected as before because the electron concentration will be the same but the movement of electrons along the wire will produce charges on a and a' which will alter the potentials V_a and $V_a{}'$ and similarly for V_b and $V_b{}'$. Assuming that equilibrium has been reached along the thermal gradients so that no e.m.f. need be considered except between metals as before we should have for the work of transfer of electrons around the circuit

$$Ne\, dE = \left(\frac{\partial \overline{F}_a}{\partial T} - \frac{\partial \overline{F}_b}{\partial T} + Ne \frac{\partial V_b}{\partial T} - Ne \frac{\partial V_a}{\partial T} \right) dT \qquad (16)$$

If $\dfrac{\partial V_b}{\partial t} - \dfrac{\partial V_a}{\partial t}$ is different from zero equation (12) does not hold. Furthermore it should be emphasized that no conclusions may be drawn as to the location of the e.m.f.'s in the thermocouple because of our lack of knowledge of the conditions for equilibrium along a gradient. The second law if it applies at all gives us information only as to the net e.m.f. of the whole circuit.

One more point needs to be emphasized. It has been commonly stated the thermoelectric power is the temperature coefficient of the contact difference of potential. If equation (12) is true then the thermoelectric power is the temperature coefficient of the "true" contact difference of potential which cannot be measured experimentally. If equation (12) is not true then there is no relation apparent between thermoelectric power and the contact potential.

[3] Compton (20) attempted to determine the relative charge at two ends of a wire in a thermal gradient and got a very large effect. It is likely that the adsorption of gas on the surface of the metal affected the results here as in most measurements of this kind.

REFERENCES

(1) MILLIKAN, R. A.: Phys. Rev., **18**, 236–244 (1921).
(2) WELCH, G. B.: Proc. Nat'l. Acad., **13**, 111–113, 1927, gives a very complete
    list of references of experimental work on the photoelectric effect.
(3) SCHOTTKY, W.: Jahrbuch der Radioactivitat, **12**, 147–205 (1915); Zeit. f.
    Physik., **14**, 63–106 (1923); Ibid., **34**, 645–675 (1925); Phys. Zeit., **20**,
    220–228 (1919); Ann. d. Physik., **62**, 113–155 (1920).
(4) LANGMUIR, J.: Trans. Am. Electrochem. Soc., **29**, 125–182 (1916).
    BRIDGEMAN, P. W.: Phys. Rev., **14**, 306–347 (1919); Ibid., **27**, 173–180 (1926).
(5) RICHARDSON, O. W.: The Emission of Electricity from Hot Bodies, Long-
    mans Green and Co., New York (1916).
(6) DUSHMAN, S.: Phys. Rev., **21**, 623–636 (1923).
    DUSHMAN AND OTHERS: Phys. Rev., **25**, 338–360 (1925).
(7) DAVISON AND GERMER: Phys. Rev., **20**, 300–330 (1922).
    DAVISON, C.: Phil. Mag., **47**, 544–549 (1924).
    HALL, E. H.: Proc. Nat'l Acad., **13**, 43–46 (1927).
(8) WARNER, A. H.: Proc. Nat'l Acad., **13**, 56–60 (1927).
(9) BORN, M.: Problems of Atomic Dynamics Mass. Inst. Technology (1926).
    MADELUNG, E.: Phys. Zeit., **19**, 524–533 (1918).
    WEIGLE, J. J.: Phys. Rev., **25**, 246 (1925).
(10) MILLIKAN AND EYRING: Phys. Rev., **27**, 51–67 (1926).
(11) LORENTZ, H. A.: Vorträge uber kinetische Theorie, p. 151, Teubner Berlin
    (1914).
(12) LEWIS AND RANDALL: Thermodynamics, Chaps. XVI, XVIII, McGraw Hill
    Co., New York (1923).
(13) GIBBS, J. W.: Sci. Papers, p. 332, Longmans Green and Co., New York
    (1906).
(14) KELVIN, LORD: Phil. Mag., **46**, 82–120 (1898).
(15) KAZDA, C. B.: Phys. Rev., **26**, 643–654 (1925).
(16) LEWIS AND RANDALL: Thermodynamics, p. 401, McGraw Hill Co., New York
    (1923).
(17) EASTMAN, LEWIS AND RODEBUSH: Proc. Nat'l Acad., **4**, 25–29 (1918).
(18) RODEBUSH, W. H.: Treatise on Physical Chemistry, p. 1196, Van Nostrand
    Co., New York (1924).
(19) LATIMER, W. M.: J. Am. Chem. Soc., **44**, 2136–2148 (1922).
(20) COMPTON, K. T.: Phys. Rev., **7**, 209–214 (1916).
(21) CHAPMAN, J. J.: Am. Chem. Soc., **48**, 2577–2589 (1926).
    EASTMAN, E. D.: Ibid., **48**, 1482–1493 (1926).
(22) COMPTON AND MOHLER: Bulletin, Nat'l Research Council, No. 48, Washing-
    ton (1924).

THE ACTIVITY COEFFICIENT OF GASES IN AQUEOUS SALT SOLUTIONS

MERLE RANDALL AND CRAWFORD FAIRBANKS FAILEY

Chemical Laboratory, University of California, Berkeley, California

GENERAL INTRODUCTION

In this and the following papers, we shall review the data for the effect of electrolytes upon the thermodynamic properties of aqueous solutions of non-electrolytes and the undissociated portion of weak electrolytes. It will be shown that the undissociated part of a moderately strong electrolyte behaves in just the same way as do typical non-polar substances.

In the development of the theory of strong electrolytes, we have used the term "activity coefficient" to express the number by which the molality of a substance must be multiplied to give the measured activity or thermodynamic behavior of that substance. The deviation of the activity coefficient from unity measures the deviation of the behavior of the substance from the laws of the perfect solution. Because the deviations of the ions, into which the strong electrolytes were assumed to be dissociated, were large, this field has received a large amount of detailed study.

Even for aqueous solutions of single non-electrolytes the experimental determinations of the activity coefficients are meager and not always concordant. Lewis and Randall (1) reviewed some of the older freezing point and vapor pressure data, and came to the conclusion that the divergence functions[1] h and j were of such a nature that h/m or j/m assumed a constant value in dilute solutions. Previously it had been assumed that moderately concentrated aqueous solutions of most non-electrolytes

---

[1] $j = 1 - \dfrac{\theta}{\lambda m}$; $h = 1 + \dfrac{\ln a_1}{r} = 1 + \dfrac{55.51 \ln a_1}{m}$ where θ = freezing point depression; m = molality; λ = 1.858; a_1 = activity of water, and r = mol fraction solute/mol fraction water.

obey Henry's Law, but Lewis and Randall calculated the activity coefficients, $\gamma = a_2/m$, of several substances, and found in aqueous solutions of glycerine as dilute as 0.1 M, $\gamma = 1.006$, and in 5 M solutions, $\gamma = 1.348$.

Various authors have considered the effect of electrolytes in lowering the solubilities of non-electrolytes, and particularly the solubilities of gases, in water.

Setschenow (2) proposed the empirical formula $S = S°e_-{}^{kc}$, which is equivalent to $kc = ln\ (S°/S)$, where k is a constant for a given salt, c is the salt concentration in mols per liter, $S°$ is the solubility of gas in pure water under standard conditions, and S is the solubility in the salt solutions.

Jahn from the measurements of Gordon (3) gave the formula, $(S° - S)/c^{\frac{1}{3}} = k$, which is also empirical. Rothmund (4) gives a simplification of Setschenow's equation, namely $(S° - S)/S° = kc$; Euler (5) and Geffcken (6) relate the "salting out" to an increase of the internal pressure of the solution caused by electrolytes. Hildebrand (7) shows that as the compressibilities of aqueous solutions increase the "salting out" diminishes.

Finally, Debye and McAulay (8) have considered the activity coefficients of non-electrolytes in the presence of electrolytes. They showed that the deviations of a_2/N_2 from unity, due to the effect of the non-electrolyte in lowering the dielectric constant of water, are given as a first approximation by the equation:

$$\ln a_2/N_2 = \alpha\ n'\ \frac{\Sigma\ \nu_i\ z_i{}^2\ e^2}{2\ D°\ r\ k\ T},\tag{1}$$

where a_2 is the activity of non-electrolyte, N_2 the mol fraction of non-electrolyte, n' the number of molecules of salt per cm^3 of solution, ν_i the number of ions of the ith kind in each salt molecule, z_i the valence of an ion of the ith kind, e the charge of an electron, $D°$ the dielectric constant of water, r the mean ionic radius, k the Boltzmann gas constant, and T the absolute temperature. The constant α is defined by the relation, $D = D°$ $(1 - \alpha n)$ where D is the dielectric constant of a solution of the non-electrolyte, and n is the number of molecules of non-electrolyte per cm^3.

Preliminary to a study of the weak electrolytes we wish to examine the numerous data with reference to this relation.

It is not practical at the present time to treat the freezing point, vapor pressure, and osmotic pressure data, for these methods give the activity of the water and we are not in a position fully to separate the effect of the electrolyte and non-electrolyte upon its activity. Also the non-electrolyte will have an effect on the activity of the electrolyte.

ACTIVITY COEFFICIENT OF GASES

The determination of the solubility of a gas is difficult; for example, it has been shown by Cady, Elsey and Berger (9) that an error of 100 per cent may be made if the liquid is violently shaken, as is usually done.[2]

We have examined all of the apparently reliable data and have summarized in the following tables the results of our study.

If we consider, for example, the reaction, $O_2(g) = O_2(aq)$, then the stoichiometrical equilibrium constant is $K_m = m/P$, where P is the pressure in atmospheres and m is the molality. The equilibrium constant is $K = a_2(aq)/a_2(g)$. For gases dissolved in pure water, $a_2(aq) = \gamma m = m$ as the solutions are so dilute that γ, the activity coefficient, is unity, and for gases at moderate pressures, within the limits of the experimental error, $a_2(g) = P$. Hence in pure water we may take $K_m = K$.

In any solution, $K = a_2(aq)/a_2(g) = \gamma m/P = \gamma K_m$. Thus, to obtain γ, for oxygen dissolved in a salt solution, we need only to know K_m in pure water, which is equal to K and also K_m in the salt solution, whence,

$$\gamma \text{ (in salt solution)} = K/K_m \text{ (in salt solution)} \qquad (2)$$

The activity of the gas has been taken equal to the partial pressure in atmospheres. This is not always justified, for in case of carbon dioxide at one atmosphere $a(g)/P = 0.99$. But since a series of measurements is made at approximately the same pressure the error of this assumption cancels out.

---

[2] The authors studied the solubility of helium in water and ascribed the supersaturation to the effect of hammering and of small bubbles.

The concentration in the liquid phase is always expressed as mols per 1000 grams water. In almost all cases it was necessary to transform the data from mols per liter to mols per 1000 grams water. This transformation was made by assuming (approximately) that $(m/c) = 1 + k\,c'$, where m is the molality of the salt or gas, c is the concentration of salt or gas in mols per liter, c' the concentration of the salt in mols per liter, and k a constant which varies from salt to salt, but is the same for solutions of one salt of varying concentration.

Most of these measurements (10) were made by shaking the salt solution in a flask connected by means of a metal capillary to a manometer, measuring the change in pressure and volume, and reducing the volume of absorbed gas to standard conditions. Abegg and Riesenfeld used a dynamic method with ammonia. McLauchlan saturated his solutions with hydrogen sulfide and then titrated them.

The data are summarized in the tables. The first column gives the salt used. The second gives the ionic strength μ, which equals the molality for uni-uni, is 3 times the molality for uni-bi, 4 times the molality for bi-bi, 9 times the molality for uni-tri and 15 times the molality for bi-trivalent salts.[3] The third column gives the activity coefficient of the gas dissolved in the salt solution, and the fourth the quotient, $(\log \gamma)/\mu$.

DISCUSSION

As previously mentioned[2] (9), there may be relatively large absolute errors in the solubility measurements. But the relative error in any series is not great as is shown by the agreement between different authors in a few cases.

We have plotted, in figures 1 and 2, the values of the quotients of $\log \gamma$ by the ionic strengths against the square roots of the ionic strengths for those concentrations below ionic strength of 4 M. The values at 15° are shown by dotted, those at 20° by broken, and those at 25° by solid, lines. In many cases the values

[3] In the case of added non-electrolytes the molality is used instead of the ionic strength.
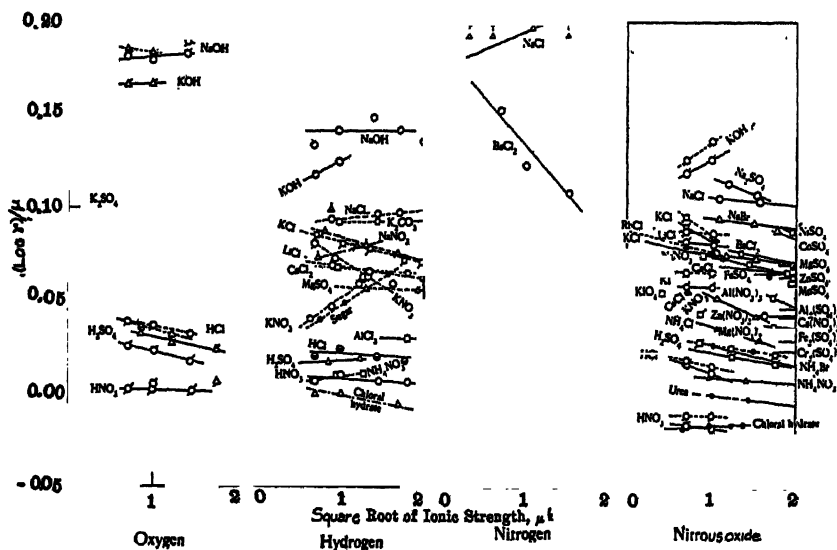
Fig. 1. Salting-out Effect of Salts on Gases
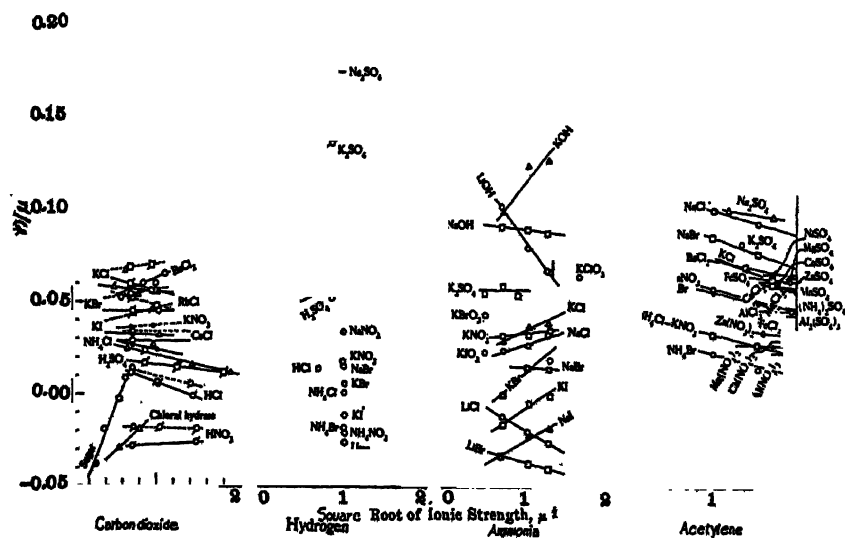
Fig. 2. Salting-out Effect of Salts on Gases

of $(\log \gamma)/\mu$ are far from constant, increasing or decreasing as the salt concentration is increased. For the most part, however, the variations lie within the probable experimental error, and are of

TABLE 1

Activity coefficient of oxygen^d in aqueous solutions*

Solubility in water $= 0.00162\ M$ at 15°—0.00138 M at 25°

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
HCl, 15°	0.505	1.046	0.0386	NaOH, 25°	0.500	1.232	0.1812
	1.019	1.090	0.0367		1.000	1.510	0.1790
	2.080	1.167	0.0322		2.000	2.316	0.1824
HCl, 25°	0.505	1.030	0.0253	NaCl, 15°	0.505	1.167	0.1330
	1.019	1.055	0.0229		1.018	1.370	0.1340
	2.080	1.108	0.0173		2.072	1.921	0.1368
H₂SO₄, 15°	0.758	1.065	0.0361	NaCl, 25°	0.505	1.167	0.1328
	1.527	1.117	0.0314		1.018	1.356	0.1299
	3.108	1.046	0.0063		2.072	1.878	0.1320
H₂SO₄, 25°	0.756	1.058	0.0324	KOH, 15°	0.503	1.239	0.1851
	1.527	1.100	0.0271		1.013	1.532	0.1829
	3.108	1.185	0.0237				
HNO₃, 15°	0.508	1.028	0.0023	KOH, 25°	0.503	1.212	0.1660
	1.030	1.049	0.0020		1.013	1.474	0.1663
	2.140	1.077	0.0015	K₂SO₄, 15°	0.760	1.210	0.1081
HNO₃, 25°	0.508	1.003	0.0025		1.563	1.470	0.1070
	1.030	1.013	0.0054	K₂SO₄, 25°	0.760	1.194	0.1006
	2.140	1.013	0.0026		1.563	1.426	0.0985
NaOH, 15°	0.500	1.260	0.2008				
	1.000	1.571	0.1962				
	2.000	2.388	0.1890				

* The letter references in these tables correspond with the citations under reference (10).

a random character. Although we have drawn the curves in the figures with definite slopes, we would probably be justified in taking the values of $(\log \gamma)/\mu$ to be constants, and use the arithmetical mean of the values given in tables 1 to 8.

TABLE 2

Activity coefficient of hydrogen in aqueous solutions

Solubility in water = 0.000886 M at 15°—0.000863 M at 25°C.

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
HCl,[d] 25°	0.505	1.024	0.0204	NaNO₃,[h] 15°	0.694	1.150	0.0875
	1.019	1.058	0.0240		1.477	1.315	0.0805
	2.080	1.101	0.0201		2.882	1.650	0.0755
	3.186	1.140	0.0179		4.485	2.062	0.0701
					6.770	2.745	0.0647
H₂SO₄,[d] 25°	0.757	1.030	0.0169				
	1.527	1.070	0.0193	NaNO₃,° 20°	0.542	1.095	0.0728
	3.110	1.140	0.0183		0.844	1.212	0.0990
					1.739	1.377	0.0800
HNO₃,[d] 25°	0.508	1.008	0.0069				
	1.030	1.024	0.0100	KOH,[d] 25°	0.504	1.146	0.118
	2.140	1.035	0.0069		1.013	1.338	0.125
	3.294	1.052	0.0066				
				KCl,[h] 15°	0.534	1.111	0.0855
AlCl₃,[h] 15°	3.290	1.252	0.0297		1.086	1.224	0.0809
	6.250	1.522	0.0292		1.850	1.397	0.0785
	11.23	2.074	0.0282		3.171	1.692	0.0720
	19.74	3.309	0.0263		3.945	1.902	0.0708
MgSO₄,[h] 15°	1.732	1.255	0.0570	KNO₃,[h] 15°	0.492	1.096	0.0809
	3.744	1.625	0.0563		0.912	1.164	0.0723
	6.520	2.363	0.0572		1.643	1.270	0.0632
	10.00	3.774	0.0577		1.962	1.333	0.0635
					2.685	1.444	0.0594
CaCl₂,[h] 15°	0.963	1.163	0.0681				
	1.734	1.299	0.0655	KNO₃,° 20°	0.413	1.039	0.0403
	3.366	1.655	0.0650		0.829	1.094	0.0471
	5.480	2.244	0.0640		1.540	1.236	0.0596
	8.870	3.628	0.0630				
				K₂CO₃,[h] 15°	1.031	1.245	0.0923
LiCl,[h] 15°	0.848	1.145	0.0694		2.103	1.567	0.0928
	1.861	1.329	0.0664		4.264	2.400	0.0890
	3.999	1.776	0.0624		6.804	3.875	0.0865
					14.37	10.70	0.0716
NaOH,[d] 25°	0.500	1.167	0.134	NH₄NO₃,° 20°	0.436	1.007	0.0070
	1.000	1.386	0.142		0.907	1.021	0.0100
	2.000	1.986	0.149		1.631	1.041	0.0107
	3.000	2.675	0.142				
	4.000	3.502	0.136	Chloral hydrate,° 20°	0.504	1.000	0.000
					1.030	1.000	0.000
NaCl,[h] 15°	0.837	1.199	0.0942		2.845	0.967	−0.0051
	2.125	1.607	0.0970		6.000	0.872	−0.0099
	2.941	1.939	0.0978	Sugar,[h] 15°	0.594	1.056	0.0398
	5.230	2.915	0.0889		1.263	1.153	0.0490
					2.481	1.446	0.0646

In making the plots of figures 1 and 2 we have used the square roots of the ionic strengths as the abscissae. This was done for convenience in evenly distributing the points, for since the values of $(\log \gamma)/\mu$ seem to be practically constant, even at very high ionic strengths (see table 8), the function used does not matter. But it is interesting that many of the curves in the figures are straight lines.

There seems to be a definite temperature coefficient, the value of $(\log \gamma)/\mu$ being slightly higher the lower the temperature in most cases.

TABLE 3

Activity coefficient of nitrogen in aqueous solutions at 25°[d]

Solubility in water = 0.000641 M at 25°

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
BaCl$_2$	0.509	1.196	0.153	Urea	0.389	1.013	0.0062
	1.070	1.355	0.123		0.912	0.999	−0.0005
	2.310	1.779	0.108		1.900	1.002	0.0005
					3.330	0.958	−0.0055
NaCl	0.115	1.096	0.346				
	0.372	1.259	0.269				
	1.170	1.698	0.197				
	2.270	2.920	0.205				

The value of the activity coefficient will depend upon the units chosen to express the concentrations. For thermodynamic purposes the molality or the mol fraction is much more useful than the concentration.

The quotient of the activity coefficient by the molality will also be constant, but by taking the quotient by the ionic strength the values of the function are brought to much closer values for similar substances, e.g., acids. In this respect equation 1 is justified.

But according to equation 1, for the same non-electrolyte in solutions of different salts the constant should vary inversely as the atomic radius of the salt.

The values of the mean ionic radius of such of the salts as can be calculated from the radii of combination given by Bragg and

TABLE 4

Activity coefficient of nitrous oxide in aqueous solutions

Solubility in water = 0.0348 M at 15°—0.0218 M at 25°

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
HCl,[d] 15°	0.505	1.022	0.0188	Cr₂(SO₄)₃[g]	9.075	1.582	0.0220
	1.018	1.036	0.0151		19.41	2.620	0.0215
HCl,[d] 25°	0.505	1.021	0.0178	Al₂(SO₄)₃[g]	7.916	2.328	0.0463
	1.018	1.028	0.0118		12.73	3.805	0.0455
H₂SO₄,[d] 15°	0.757	1.050	0.0280	Al(NO₃)₃[g]	2.972	1.428	0.0520
	1.527	1.093	0.0253		6.144	1.702	0.0375
	3.108	1.165	0.0213	MgSO₄[g]	3.633	1.794	0.0698
H₂SO₄,[d] 25°	0.757	1.041	0.0433		7.280	3.290	0.0710
	1.527	1.075	0.0206	Mg(NO₃)₂[g]	3.066	1.291	0.0361
	3.108	1.128	0.0168		6.447	1.681	0.0349
HNO₃,[d] 15°	0.508	0.987	−0.0112	CaCl₂[g]	2.861	1.536	0.0652
	1.030	0.973	−0.0116		6.310	2.501	0.0631
HNO₃,[d] 25°	0.508	0.981	−0.0164	Ca(NO₃)₂[g]	4.422	1.532	0.0414
	1.030	0.958	−0.0180	BaCl₂[g]	1.898	1.395	0.0762
ZnSO₄,[g] 25°	3.850	1.758	0.0636		4.122	1.953	0.0705
	7.564	3.062	0.0642	LiCl,[d] 15°	0.505	1.107	0.0877
Zn(NO₃)₂,[g] 25°	2.616	1.291	0.0423		1.019	1.226	0.0869
	5.448	1.697	0.0421	LiCl,[d] 25°	0.505	1.100	0.0821
Cu(NO₃)₂,[g] 25°	2.127	1.456	0.0767		1.019	1.208	0.0806
	4.377	1.813	0.0590	NaCl[g]	1.178	1.333	0.1059
MnSO₄,[g] 25°	3.834	1.708	0.0606		2.426	1.782	0.1034
	8.088	2.993	0.0588		4.761	2.805	0.0940
FeSO₄,[g] 25°	2.902	1.554	0.0659	NaBr[g]	1.161	1.287	0.0944
	5.852	2.423	0.0656		2.306	1.625	0.0914
Fe₂(SO₄)₃,[g] 25°	10.31	1.975	0.0286		5.098	2.625	0.0822
	21.90	3.673	0.0257	Na₂SO₄[g]	1.409	1.444	0.1132
CoSO₄,[g] 25°	3.181	1.917	0.0888		3.006	2.098	0.1071
	6.448	3.048	0.0750	NaNO₃[g]	1.120	1.214	0.0751
NiSO₄,[g] 25°	3.784	2.141	0.0873		1.370	1.266	0.0747
	7.696	3.752	0.0746		2.334	1.480	0.0729
					3.350	1.730	0.0711
					4.894	2.115	0.0664

TABLE 4—*Continued*

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
KOH,[d] 15°	0.503	1.158	0.1266	KNO$_3$,[e] 25°	1.068	1.136	0.0518
	1.013	1.375	0.1365		2.374	1.250	0.0422
KOH,[d] 25°	0.503	1.149	0.1199	RbCl,[d] 15°	0.509	1.099	0.0806
	1.013	1.344	0.1268		1.036	1.201	0.0767
KCl,[d] 15°	0.508	1.118	0.0954	RbCl,[d] 25°	0.509	1.094	0.0766
	1.031	1.226	0.0858		1.036	1.188	0.0722
KCl,[d] 25°	0.508	1.111	0.0900	CsCl,[d] 15°	0.511	1.072	0.0591
	1.031	1.213	0.0814	CsCl,[d] 25°	0.511	1.069	0.0567
KCl,[e] 25°	0.800	1.148	0.0748				
	1.301	1.248	0.0739		1.118	1.094	0.0348
	2.112	1.406	0.0701	NH$_4$Cl,[e] 25°	2.470	1.181	0.0292
	3.570	1.708	0.0651		5.185	1.275	0.0203
	4.632	1.938	0.0620				
KBr,[d] 15°	0.510	1.095	0.0773		1.093	1.070	0.0269
	1.041	1.192	0.0732	NH$_4$Br,[e] 25°	2.341	1.136	0.0236
KBr,[d] 25°	0.510	1.086	0.0702		5.161	1.195	0.0150
	1.041	1.177	0.0680				
	1.160	1.187	0.0642		0.933	1.021	0.0096
KBr,[e] 25°	2.339	1.396	0.0619	NH$_4$NO$_3$,[e] 25°	2.045	1.038	0.0079
	4.992	1.813	0.0517		4.463	1.059	0.0055
KI,[d] 15°	0.513	1.081	0.0660		11.95	1.072	0.0025
	1.050	1.171	0.0653				
				Chloral	0.445	0.981	−0.0187
				hydrate,[e] 20°	0.942	0.964	−0.0169
KI,[d] 25°	0.513	1.071	0.0581		1.474	0.946	−0.0164
	1.050	1.149	0.0574		1.911	0.930	−0.0165
					1.016	0.998	−0.0009
KIO$_4$,[e] 25°	0.164	1.021	0.0548	Urea,[e] 25°	2.139	0.987	−0.0026
					4.955	0.942	−0.0052
					7.995	0.883	−0.0067

Bragg (11) are given in Table 9. The values are obtained by means of the formula

$$r = \Sigma \, \nu_i \, z_i^2 / \Sigma \, (\nu_i \, z_i^2 / r_i), \qquad (3)$$

where r_i is the radius of combination of an ion of the ith kind, and r is the mean atomic radius for the molecule, and ν_i and z_i have the same significance as in equation 1.

TABLE 5

Activity coefficient of carbon dioxide in aqueous solutions

Solubility in water $= 0.0478\ M$ at 15°—0.0370 M at 25°

SALT	μ	γ	$(\log\gamma)/\mu$	SALT	μ	γ	$(\log\gamma)/\mu$
HCl,[d] 15°	0.522	1.017	0.0140	KCl,[c] 25°	1.031	1.141	0.0555
	1.028	1.021	0.0070	KBr,[d] 15°	0.493	1.064	0.0546
	2.000	1.029	0.0062		0.914	1.125	0.0560
HCl,[d] 25°	0.505	1.015	0.0129	KBr,[d] 25°	0.510	1.054	0.0448
	1.019	1.014	0.0059		1.041	1.113	0.0447
	2.080	0.998	−0.0004	KI,[d] 15°	0.496	1.052	0.0444
H$_2$SO$_4$,[d] 15°	0.757	1.042	0.0236		0.923	1.104	0.0466
	1.554	1.056	0.0153	KI,[d] 25°	0.513	1.040	0.0332
	3.216	1.089	0.0116		1.050	1.082	0.0318
H$_2$SO$_4$,[d] 25°	0.757	1.031	0.0176	KNO$_3$,[d] 15°	0.503	1.043	0.0364
	1.527	1.053	0.0147		0.946	1.084	0.0370
	3.108	1.092	0.0123	KNO$_3$,[d] 25°	0.511	1.032	0.0268
HNO$_3$,[d] 15°	0.539	0.978	−0.0180		1.043	1.060	0.0243
	1.086	0.956	−0.0180	RbCl,[d] 15°	0.509	1.063	0.0521
	2.200	0.909	−0.0188		1.036	1.122	0.0483
HNO$_3$,[d] 25°	0.508	0.968	−0.0278	CsCl,[d] 15°	0.511	1.041	0.0342
	1.030	0.960	−0.0172	CsCl,[d] 25°	0.511	1.034	0.0284
	2.140	0.880	−0.0259	NH$_4$Cl,[c] 25°	0.446	1.026	0.0251
BaCl$_2$,[c] 25°	0.345	1.042	0.0518		0.947	1.060	0.0267
	0.718	1.106	0.0609		1.630	1.063	0.0163
	1.010	1.151	0.0604		2.010	1.080	0.0166
	1.239	1.204	0.0650		3.600	1.103	0.0118
KCl,[d] 15°	0.488	1.080	0.0685	Chloral hydrate,[c] 25°	0.317	0.979	−0.0290
	0.897	1.157	0.0706		0.654	0.972	−0.0188
KCl,[d] 25°	0.508	1.072	0.0595	Sugar,[c] 25°	0.078	0.993	−0.0382
	1.031	1.143	0.0563		0.157	0.993	−0.0191
KCl,[c] 25°	0.248	1.034	0.0584		0.305	0.998	−0.0029
	0.414	1.066	0.0671		0.396	1.009	0.0098
	0.626	1.081	0.0539				

TABLE 6

Activity coefficient of hydrogen sulphide in aqueous solutions at 25°[t]

Solubility in water = 0.000135 M under pressure of 1 mm. Hg at 25°

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
HCl	0.505	1.016	0.0136	KBr	1.041	1.016	0.0062
H₂SO₄	0.758	1.095	0.0519	KI	1.050	0.972	−0.0117
NaCl {	0.505	1.066	0.0550	K₂SO₄ {	0.377	1.118	0.1283
	1.018	1.160	0.0633		0.758	1.269	0.1365
NaBr	1.029	1.039	0.0161	KNO₃	1.043	1.050	0.0183
Na₂SO₄ {	0.377	1.164	0.1750	NH₄Cl	1.039	1.003	0.0013
	0.758	1.357	0.1749	NH₄Br	1.046	0.956	−0.0184
NaNO₃	1.032	1.084	0.0339	NH₄NO₃	1.064	0.950	−0.0209
KCl	1.031	1.137	0.0541	Urea	1.045	0.938	−0.0266

TABLE 7

Activity coefficient of ammonia in aqueous solutions at 25°[a]

Solubility in water = 0.00743 M under pressure of 1 mm. Hg. at 25°

SALT	μ	γ	$(\log \gamma)/\mu$	SALT	μ	γ	$(\log \gamma)/\mu$
LiOH {	0.513	1.128	0.1019	KOH {	0.516	1.139	0.1094
	1.025	1.207	0.0797		1.038	1.347	0.1246
	1.538	1.270	0.0674		1.566	1.580	0.1268
LiCl {	0.517	0.986	−0.0117	KCl {	0.520	1.036	0.0296
	1.044	0.951	−0.0208		1.056	0.093	0.0365
	1.580	0.909	−0.0262		1.605	1.155	0.0390
LiBr {	0.520	0.961	−0.0332	KClO₃	0.260	1.039	0.0638
	1.053	0.914	−0.0370				
	1.601	0.861	−0.0405	KBr {	0.523	1.006	0.0050
					1.066	1.038	0.0152
LiI {	0.523	0.928	−0.0619		1.629	1.075	0.0193
	1.066	0.856	−0.0633				
	1.629	0.771	−0.0693	KBrO₃	0.259	1.026	0.0432
NaOH {	0.513	1.114	0.0914	KI {	0.525	0.981	−0.0158
	1.025	1.239	0.0908		1.075	0.988	−0.0048
	1.538	1.363	0.0874		1.650	1.000	0.0000
NaCl {	0.517	1.029	0.0240	KIO₃	0.259	1.014	0.0231
	1.043	1.079	0.0276				
	1.573	1.128	0.0332	K₂SO₄ {	0.259	1.104	0.0553
					0.523	1.238	0.0590
NaBr {	0.520	0.999	−0.0008		0.816	1.356	0.0540
	1.054	1.036	0.0146				
	1.602	1.052	0.0137	KNO₃ {	0.523	1.037	0.0302
					1.068	1.087	0.0338
NaI {	0.523	0.961	−0.0331		1.634	1.143	0.0355
	1.066	0.947	−0.0221				
	1.629	0.935	−0.0179				

TABLE 8
Activity coefficient of acetylene in aqueous solutions at 25°

Solubility in water = 0.0421 M at 25°

SALT		μ	γ	$(\log \gamma)/\mu$	SALT		μ	γ	$(\log \gamma)/\mu$
ZnSO₄	{	3.704	1.701	0.0622	CaCl₂	{	2.938	1.518	0.0617
	{	7.536	2.878	0.0609		{	6.513	2.273	0.0547
Zn(NO₃)₂	{	2.580	1.222	0.0337	Ca(NO₃)₂	{	4.422	1.350	0.0294
	{	5.367	1.465	0.0308		{	9.696	1.812	0.0266
MnSO₄	{	3.999	1.688	0.0568	BaCl₂	{	1.936	1.366	0.0699
	{	8.212	2.869	0.0557		{	4.128	1.833	0.0637
FeSO₄	{	2.902	1.520	0.0626	NaCl	{	1.147	1.305	0.1007
	{	5.852	2.258	0.0604		{	2.372	1.662	0.0930
						{	5.012	2.666	0.0849
FeCl₃	{	5.129	1.453	0.0316					
	{	8.412	1.735	0.0284	NaBr	{	1.158	1.256	0.0854
						{	2.393	1.525	0.0765
Fe₂(SO₄)₃	{	10.31	1.979	0.0287		{	5.152	2.274	0.0692
	{	21.90	3.761	0.0262					
					NaSO₄	{	1.420	1.390	0.1007
CoSO₄	{	3.219	1.589	0.0624		{	2.978	1.934	0.0962
	{	6.528	2.484	0.0605					
					NaNO₃	{	1.121	1.160	0.0575
NiSO₄	{	2.966	1.571	0.0661		{	2.330	1.312	0.0506
	{	5.980	2.452	0.0651		{	5.072	1.641	0.0424
Cr₂(SO₄)₃	{	9.075	1.568	0.0215	KCl	{	1.956	1.361	0.0684
	{	19.41	2.565	0.0210		{	4.176	1.736	0.0573
AlCl₃	{	3.798	1.491	0.0456	KBr	{	1.139	1.158	0.0559
	{	6.642	1.946	0.0435		{	2.381	1.317	0.0502
						{	5.272	1.634	0.0404
Al₂(SO₄)₃	{	7.916	2.221	0.0437					
	{	12.73	3.562	0.0433	K₂SO₄		1.834	1.415	0.0822
					KNO₃	{	1.129	1.091	0.0334
Al(NO₃)₃	{	3.460	1.204	0.0232		{	2.376	1.160	0.0271
	{	6.576	1.401	0.0222					
					NH₄Cl	{	1.137	1.091	0.0332
MgCl₂	{	3.396	1.576	0.0581		{	2.381	1.154	0.0261
	{	6.738	2.251	0.0523		{	5.254	1.215	0.0161
MgSO₄	{	3.615	1.700	0.0637	NH₄Br	{	1.125	1.061	0.0228
	{	7.328	2.968	0.0644		{	2.372	1.078	0.0137
						{	5.323	1.104	0.0082
Mg(NO₃)₂	{	2.998	1.207	0.0272	(NH₄)₂SO₄	{	4.533	1.638	0.0472
	{	6.333	1.421	0.0240		{	8.181	2.201	0.0423

Thus the ratio of this constant for two given salts, say sodium chloride and potassium iodide, should be independent of the non-electrolyte whose activity coefficient is being measured. Or, for two different gases, say oxygen and nitrous oxide, in a solution of the same salt the ratio of the constants should be independent of

TABLE 9

Mean atomic radius of some halides ($\times 10^8$)

	H	Li	Na	K	Rb	Cs	Mg	Ca	Ba	Al	Fe^{+++}
Cl.................	0.85	1.21	1.29	1.37	1.41	1.45	1.27	1.41	1.57	1.27	1.28
Br................	0.90	1.31	2.13	1.50	1.55	1.60					
I..................	0.93	1.42	2.25	1.64	1.70	1.76					

the salt used. There seems to be a qualitative agreement with these demands in most cases, but not a quantitative one. In the case of ammonia the order of increasing salting out power (increase in the quotient) for salts of the alkali metals is that of increasing atomic radius rather than of increasing reciprocal of the atomic radius.

REFERENCES

(1) LEWIS AND RANDALL: Thermodynamics and the Free Energy of Chemical Substances, McGraw-Hill Book Co., New York (1923), pp. 275, 288.
(2) SETSCHENOW: Ann. chim. phys. [6], 25, 226 (1892).
(3) GORDON: Z. physik. Chem., 18, 8 (1895).
(4) ROTHMUND: Ibid., 33, 401 (1900).
(5) EULER: Ibid., 31, 360 (1899).
(6) GEFFCKEN: Ibid., 49, 257 (1904).
(7) HILDEBRAND: Solubility, Chemical Catalog Co., New York (1924), p. 140.
(8) DEBYE AND McAULAY: Physik. Z., 26, 22 (1925).
(9) CADY, ELSEY AND BERGER: J. Am. Chem. Soc., 44, 1456 (1922).
(10) (a) ABEGG AND RIESENFELD: Z. physik. Chem., 40, 90 (1903), NH₃.
    (b) BRAUN: Ibid., 33, 732 (1900), N₂.
    (c) FINDLAY AND SHEN: J. Chem. Soc., 101, 1461 (1912), CO₂.
    (d) GEFFCKEN: Z. physik. Chem., 49, 257 (1904), O₂, H₂, CO₂, N₂O.
    (e) KNOPP: Ibid., 48, 97 (1904), H₂, N₂O.
    (f) McLAUCHLAN: Ibid., 44, 600 (1903), H₂S.
    (g) MANCHOT, JAHRSTORFER AND ZEPTER: Z. anorg. allgem. Chem., 141, 45 (1924), N₂O, C₂H₂.
    (h) STERNER: Diss. Stuttgart (1894), H₂.
(11) BRAGG AND BRAGG: X-rays and Crystal Structure, G. Bell and Sons, London, 4th Edition (1924), p. 170.

THE ACTIVITY COEFFICIENT OF NON-ELECTROLYTES IN AQUEOUS SALT SOLUTIONS FROM SOLUBILITY MEASUREMENTS. THE SALTING-OUT ORDER OF THE IONS

MERLE RANDALL AND CRAWFORD FAIRBANKS FAILEY

University of California, Berkeley, California

In the preceding paper we studied the activity coefficient of gases dissolved in aqueous salt solutions. It was found that the quotient of the logarithm of the activity coefficient of the dissolved gas by the ionic strength, μ, of the salt was approximately a constant.

If the activity of the non-electrolyte is determined by having a solid phase in equilibrium, rather than by the pressure of a gas, the accuracy of the results should be greater. In the present paper we study the activity coefficient of iodine, phenylthiourea, and *o*-nitrobenzaldehyde.

The solubility of these substances in water is small and we may take the activity, a, of the non-electrolyte in its saturated aqueous solution equal to its molality, $m°$, and its activity coefficient, γ_u, is, therefore, unity. In the presence of the salt, since the solid phase is present, $a_2 = m\gamma = $ constant. Hence, the activity coefficient in any salt solution is given by the expression, $\gamma_u = m°/m$, where m is the molality of the non-electrolyte in the salt solution.

Linderström-Lang (1) in a recent paper has investigated the effect of a large number of salts on the solubilities of hydroquinone, quinone, succinic acid, and boric acid. He calculated the quantity $(\log(S°/S))/c$ where $S°$ is the solubility in pure water of the substance investigated, expressed as mols per liter, S its solubility in the salt solution, and c the equivalent concentration of the salt. This quantity $(\log(S°/S))/c$ was found to be approximately constant for a given saturating substance and a

single salt at different concentrations, increasing with concentrations of salt in some cases, in others decreasing. We do not have the necessary density determinations to enable us to calculate $(\log \gamma_u)/\mu$. From approximate densities we estimate that this quantity will also be approximately constant, its variation being in some cases greater, in some less, than Linderström-Lang's expression.

TABLE 1

Activity coefficient of iodine[2] in aqueous salt solutions[1]

Solubility in water = 0.001321 M at 25°—0.001815 M at 35°

SALT	μ	γ_u	$\dfrac{\log \gamma_u}{\mu}$	SALT	μ	γ_u	$\dfrac{\log \gamma_u}{\mu}$
Na₂SO₄, 25°	1.267	1.376	0.1094	NaNO₃, 25°	6.455	2.204	0.0532
	1.997	1.653	0.1093		7.787	2.613	0.0536
	4.095	2.784	0.1086		9.302	3.161	0.0537
	5.166	3.628	0.1083				
Na₂SO₄, 35°	0.224	1.028	0.0535	NaNO₃, 35°	0.2376	1.000	0.0000
	0.499	1.116	0.0955		0.4241	1.022	0.0223
	1.095	1.298	0.1034		0.8651	1.083	0.0400
	2.122	1.619	0.0986		2.222	1.220	0.0389
	3.666	2.459	0.1066		4.818	1.627	0.0439
	5.646	3.897	0.1046		7.898	2.287	0.0455
	7.155	5.657	0.1052		11.81	3.270	0.0436
	9.843	10.72	0.1047				
	10.45	12.67	0.1056				
NaNO₃, 25°	0.736	1.072	0.0410	NaH₂PO₄, 25°	0.5897	1.227	0.1507
	1.469	1.178	0.0484		1.218	1.478	0.1393
	2.120	1.287	0.0517		1.883	1.763	0.1308
	4.063	1.641	0.0529		2.597	2.036	0.1189
	4.991	1.904	0.0560		3.355	2.665	0.1269
					5.044	5.535	0.1473
					5.987	13.22	0.1873

[1] The numerical citations in these tables are to the corresponding references.

The measurements of Carter (2) on the solubility of iodine, of Biltz (3) and of Rothmund (4) on the solubility of phenylthiourea, and of Goldschmidt and Sunde (5) on the solubility of o-nitrobenzaldehyde, have been converted to molalities, and the activity coefficients calculated are given in tables 1 to 3. The first columns give the salt, the second the ionic strength of the salt, the third the ratio of the solubility in pure water to that in

TABLE 2

Activity coefficient of phenylthiourea in aqueous salt solutions at 20°

Solubility in water = 0.01394 M at 20°

SALT	μ	γ_u	$(\log \gamma_u)/\mu$	SALT	μ	γ_u	$(\log \gamma_u)/\mu$
AlCl₃³	0.250	1.076	0.1272	Na₂SO₄⁴	0.189	1.054	0.1204
	0.500	1.087	0.0724		0.378	1.123	0.1333
	1.003	1.155	0.0624		0.758	1.272	0.1378
	2.014	1.296	0.0559		1.532	1.646	0.1412
MgSO₄⁴	0.250	1.040	0.0680	NaNO₃⁴	0.126	0.996	−0.0134
	0.500	1.091	0.0756		0.252	1.004	0.0067
	1.000	1.208	0.0821		0.508	1.030	0.0251
	2.000	1.478	0.0849		1.032	1.074	0.0300
BaCl₂²	0.189	1.062	0.1381	KCl³	0.125	1.036	0.1232
	0.376	1.076	0.0845		0.252	1.087	0.1436
	0.756	1.132	0.0712		0.508	1.127	0.1021
	1.523	1.315	0.0780		1.031	1.282	0.1046
	3.093	1.591	0.0652	KClO₃³	0.126	1.000	0.0000
LiNO₃⁴	0.125	0.996	−0.0136		0.253	1.012	0.0205
	0.252	0.991	−0.0154		0.513	1.036	0.0300
	0.507	0.986	−0.0120	KBr³	0.126	1.027	0.0921
	1.031	0.984	−0.0067		0.253	1.034	0.0573
NaCl³	0.125	1.048	0.1632		0.510	1.067	0.0552
	0.251	1.082	0.1362		1.041	1.139	0.0542
	0.505	1.161	0.1283	KI³	0.126	0.982	−0.0626
	1.018	1.367	0.1333		0.253	0.952	−0.0845
NaClO₃³	0.126	1.009	0.0309		0.513	0.950	−0.0434
	0.252	1.012	0.0206		1.050	0.909	−0.0394
	0.510	1.047	0.0392	K₂SO₄⁴	0.189	1.049	0.1100
	1.039	1.098	0.0391		0.377	1.111	0.1212
NaClO₄³	0.126	0.980	−0.0698		0.758	1.242	0.1241
	0.253	0.980	−0.0347		1.532	1.565	0.1269
	0.512	1.002	0.0017	KNO₃⁴	0.126	0.999	−0.0038
	1.048	1.059	0.0237		0.253	0.996	−0.0067
NaI³	0.126	0.992	−0.0277		0.511	1.010	0.0084
	0.253	0.981	−0.0328		1.043	1.043	0.0175
	0.510	0.956	−0.0382				
	1.041	0.959	−0.0174				

TABLE 2—*Continued*

SALT	μ	γ_u	$(\log \gamma_u)/\mu$	SALT	μ	γ_u	$(\log \gamma_u)/\mu$
KAc[4]	0.126	1.033	0.1110	CsNO₃[3]	0.126	0.953	−0.1658
	0.254	1.061	0.1010		0.253	0.924	−0.1355
	0.515	1.115	0.0918		0.514	0.888	−0.1003
	1.060	1.224	0.0828				
RbNO₃[3]	0.126	0.975	−0.0873	NH₄NO₃[4]	0.126	0.976	−0.0833
	0.253	0.955	−0.0790		0.254	0.952	−0.0842
	0.512	0.946	−0.0470		0.516	0.930	−0.0610
	1.047	0.936	−0.0274		1.064	0.879	−0.0526

the salt solution, and the fourth the quotient of logarithm of activity coefficient by ionic strength.

The results are also shown in figure 1, in which the quotients in column 4 are plotted against the square roots of the ionic strengths.

The measurements with iodine are the most accurate. With sodium sulfate and nitrate the quotient $(\log \gamma_u)/\mu$ is nearly con-

TABLE 3

Activity coefficient of o-nitrobenzaldehyde[5] in aqueous salt solutions at 25°

Solubility in water = 0.01534 M at 25°

SALT	μ	γ_u	$(\log \gamma_u)/\mu$	SALT	μ	γ_u	$(\log \gamma_u)/\mu$
HCl	0.505	0.959	−0.0362	NaNO₃	1.032	0.718	−0.139.
	1.019	0.920	−0.0355		2.130	0.680	−0.079
	2.080	0.838	−0.0369	KCl	1.031	1.108	0.0400
HNO₃	0.508	0.711	−0.292		2.120	1.144	0.0276
	1.030	0.599	−0.216	KNO₃	0.511	0.725	−0.274
NaCl	1.018	1.205	0.0795		1.043	0.694	−0.152
	2.076	1.604	0.0990		2.170	0.625	−0.094
NaNO₃	0.508	0.756	−0.239				

stant, as was the case with gases. The ionic strength of the sodium dihydrogen phosphate (taken equal to its molality) is not simply determined and the variation from constancy is probably due to this cause. The results with the phenylthiourea and o-nitrobenzaldehyde are not very accurate, but they show the same

qualitative agreement with the demands of equation 1 of the previous paper as was shown by the gases.

The measurements with iodine at 25° and 35° also show that the quotient is somewhat larger (salting out effect is larger) at the lower temperature.

THE SALTING-OUT ORDER OF THE IONS

The use of the quotient of the logarithm of the activity coefficient by the ionic strength rather than the molality or the
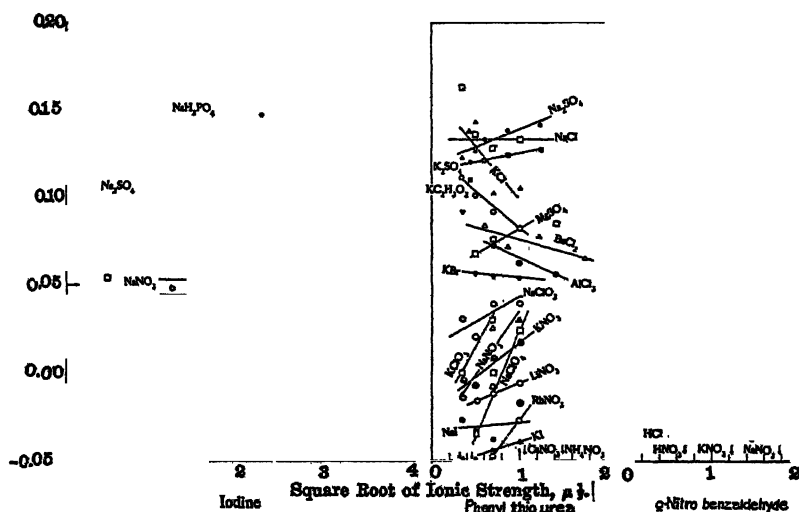
FIG. 1. SALTING-OUT EFFECT OF SALTS ON NON-ELECTROLYTES

equivalent molality, as a measure of the salting-out effect, gives us a new basis for arranging the salting-out series of the ions. The valence of the ion in our new series is of little importance. For example, barium chloride and potassium chloride have about the same effect.

It is impossible to fix an absolute order on the basis of the rather approximate data considered, but we may place the negative ions in the order of their decreasing salting-out effect about as follows: hydroxide; sulfate and carbonate; chlorate, bromate, chloride, acetate, iodate and perhalide; bromide and iodide;

nitrate. For the positive ions the order is approximately as follows: sodium; potassium; lithium, barium, rubidium, calcium, nickel, cobalt, magnesium, ferrous, zinc, cesium, manganous, aluminium, ferric, and chromic; ammonium; hydrogen.

Referring to figure 1, and to figures 1 and 2 of the preceding article, it is interesting to note that the spread of the values is different for the different non-electrolytes. That is, the difference between the effect of sodium hydroxide and nitric acid is large in some cases (cf. oxygen), and in some cases small (cf. nitrous oxide). Also the relative position with reference to a negligible effect, zero value of $(\log \gamma_u)/\mu$, is variable. Thus, all the values in the case of oxygen are positive, while many of those for phenylthiourea are negative. We have not discovered any relation between these observations and properties of the non-electrolytes.

In the case of o-nitrobenzaldehyde the effect of nitric acid and the nitrates was very large negatively. This may be due to the nitro group in the non-electrolyte.

REFERENCES

(1) LINDERSTRÖM-LANG: Compt. rendus trav. lab. Carlsberg, **15**, no. 4 (1924).
(2) CARTER: J. Chem. Soc., **127**, 2861 (1925).
(3) BILTZ: Z. physik. Chem., **43**, 44 (1903).
(4) ROTHMUND: Ibid., **33**, 401 (1900).
(5) GOLDSCHMIDT AND SUNDE: Z. physik. Chem., **56**, 1 (1906).

THE ACTIVITY COEFFICIENT OF THE UNDISSOCIATED
PART OF WEAK ELECTROLYTES

MERLE RANDALL AND CRAWFORD F. FAILEY

University of California, Berkeley, California

In the preceding two papers it was shown that the quotient of the logarithm of the activity coefficient of gases and non-electrolytes by the ionic strength of an added salt was approximately constant. This paper will show that this law is also true for the undissociated part of a weak electrolyte, when very simple assumptions are made regarding the activity coefficient of the ions.

The effect of salts upon the activity of weak electrolytes has long been a subject of controversy. Various authors, notably Arrhenius (1) have advanced the view that the dissociation of weak acids is increased by salts. McBain and Coleman (2) have objected to this view. They concluded that there was no experimental evidence for the supposed increase in the dissociation constant of weak acids and bases in the presence of neutral salts.

McBain and Kam (3) determined the vapor pressure of acetic acid over solutions containing added salt and concluded "the undissociated acid must be regarded as exhibiting enhanced chemical potential in the presence of such salts." We have recalculated their results in table 5, which shows that the effect of salts on the undissociated part of acetic acid is approximately the same as that upon non-electrolytes.

Brönsted (4) points out that the form of the solubility curves of weak acids in salt solutions are to be explained as due to the rapid decrease in the activity coefficient of the ions at low salt concentrations and the increase in γ_u (undissociated) at high salt concentration. Rördam (5) measured the solubility of benzoic, o-toluylic, and o-nitrobenzoic acids in salt solutions. He calculated the activity coefficient of the undissociated molecules and found it to increase with the salt concentration.

The activities of the ions of sodium, and potassium acetates in pure aqueous solutions have been shown by Randall, McBain and White (6) to closely approximate the activity coefficient of the corresponding chlorides. We shall in the following make the approximate assumption that the activity coefficient of the dissociated part of any monobasic acid at small molalities, in salt solutions of varying ionic strengths is equal to the activity coefficient of hydrochloric acid in the same or similar salt solution. Further we shall assume in all solutions of sodium salts the values of Harned (7) for the activity coefficient of 0.01 M hydrochloric acid in solutions of sodium chloride referred to $\gamma_{\pm} = 0.795$ in 0.1 M pure hydrochloric acid. We shall take the activity coefficient in solutions of potassium salts as equal to that found by Harned for the activity coefficient of 0.01 M hydrochloric acid in a solution of potassium chloride having the same ionic strength as the solution in which our weak acid is dissolved. The activity coefficient in solutions of barium salts is taken from the work of Randall and Breckenridge (8). These assumptions are not fully justified, but their use in the absence of experiments on very dilute hydrochloric acid in the presence of these salts, does not greatly affect our calculated values for the activity coefficient of the undissociated molecules.

We shall take the dissociation constant, K, of the acid as that determined from conductivity data, corrected in case K is large, for changes in activity coefficients and ionic mobilities by the method of Sherrill and Noyes (9).

ACTIVITY COEFFICIENT OF THE UNDISSOCIATED MOLECULES FROM SOLUBILITY MEASUREMENTS

When a solution is saturated with a weak acid the activity of the acid is fixed; $a_2 = \text{constant} = m_+ m_- \gamma_{\pm}^2$, where m_+ and m_- are the molalities of the hydrogen and acid ions, and since $m_+ = m_-$, we have $m_{\pm} \gamma_{\pm}$ also constant. By dividing this constant quantity (calculated from K, $\gamma_{\pm}$, and the solubility in pure water by successive approximations) by the value of the activity coefficient of hydrochloric acid in a salt solution of ionic strength equal to that in which the solubility determination was made, we obtain the

molality of hydrogen ion m_+, assuming always that $\gamma_\pm$ may be obtained by the principle of ionic strength. Subtracting m_+ from the solubility m, gives us the molality of undissociated acid m_u. To obtain the activity coefficient of the undissociated molecules of the acid, we divide the molality of the undissociated molecules in pure water, $m_u°$, by their molality in the salt solution: $\gamma_u = m_u°/m_u$. We then find that for a given acid in a solution of a given salt at various ionic strengths: (log γ_u)/μ(salt) = approximately constant.

Calculations of $m_u°$ *and* $m_+\gamma_\pm$ *for pure saturated solutions of weak acids in water.* Given m, the solubility in water, K, the dissociation constant, and a plot of the activity coefficient of hydrochloric acid against ionic strength in pure solution (10), we can obtain $m_u°$ and $m_+\gamma_\pm$ by successive approximations. The molality of undissociated molecules, $m_u°$ is equal to the solubility in water less m_+, the molality of the hydrogen ion. Setting the activity coefficient of undissociated molecules equal to unity, $m_u° = m - m_+$, also $m_+ = m_-$. Therefore $(m_+\gamma_\pm)^2 = K\ (m - m_+)$. Solving for m_+ we introduce a new value of $\gamma_\pm$ corresponding to the ionic strength found, and repeat the calculation. Thus we obtain m_+, $m_+\gamma_\pm$, and $m_u°$.

The data of Hoffmann and Langebeck (11) and of Rördam (5) on the solubility of benzoic, o-toluylic, salicylic, and o-nitrobenzoic acids are given in tables 1 to 4. In each table the first column gives the added salt, the second the ionic strength of the added salt, the third the molality of the dissolved weak acid, the fourth the activity coefficient of hydrochloric acid at the ionic strength of the solution, which is the sum of the third and fifth columns. The fifth column is the quotient of $m_+\gamma_\pm$, which in the saturated solution of a given acid is a constant, by $\gamma_\pm$. The sixth column is the molality of the undissociated acid or the difference between Cols. 3 and 5. The seventh column gives the activity coefficient of the undissociated part of the acids, and the last the quotient of the logarithm of the undissociated part by the ionic strength of the added salt. The ionic strength of the added salt, rather than the ionic strength of the solution, is used in obtaining this last quotient, because the salting-out effect of acids
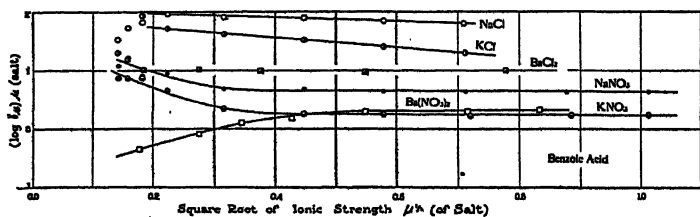
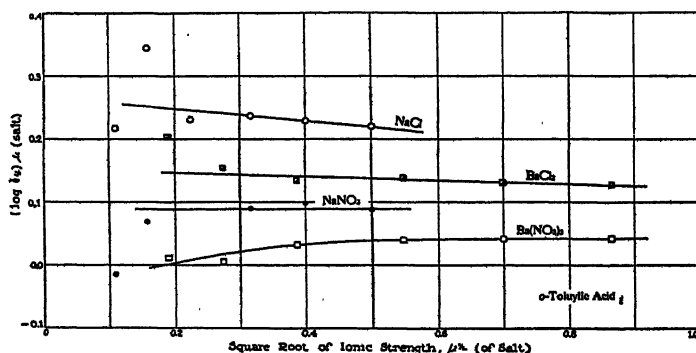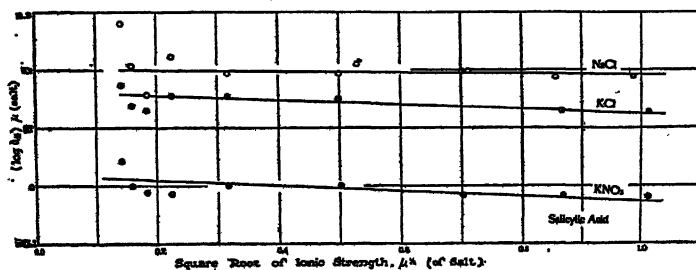FIG. 1. ACTIVITY COEFFICIENT OF UNDISSOCIATED BENZOIC ACID

FIG. 2. ACTIVITY COEFFICIENT OF UNDISSOCIATED o-TOLUYLIC ACID

FIG. 3. ACTIVITY COEFFICIENT OF UNDISSOCIATED SALICYLIC ACID
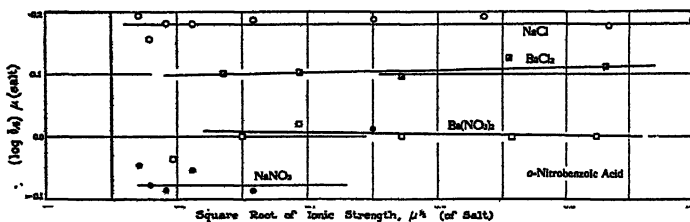

FIG. 4. ACTIVITY COEFFICIENT OF UNDISSOCIATED o-NITROBENZOIC ACID

TABLE 1

Activity coefficient of the undissociated molecules of benzoic acid in salt solutions at 25°*

$K_{298.1} = 6.61 \times 10^{-5}$, Solubility in water $= 0.02793$, $m_+ \gamma_\pm = 0.001326$, $m_u{}^\circ = 0.02655$

SALT	μ	m	$\gamma_\pm$	m_+	m_u	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$
NaCl[11]	0.01982	0.02790	0.864	0.00153	0.02637	1.007	0.153
	0.02502	0.02783	0.862	0.00154	0.02629	1.010	0.173
	0.03320	0.02775	0.849	0.00156	0.02619	1.014	0.182
	0.05010	0.02755	0.829	0.00160	0.02595	1.023	0.197
	0.1002	0.02708	0.789	0.00168	0.02540	1.045	0.191
	0.2009	0.02607	0.755	0.00176	0.02431	1.092	0.190
	0.3356	0.02481	0.737	0.00180	0.02301	1.154	0.185
	0.5022	0.02338	0.733	0.00181	0.02157	1.231	0.180
NaNO₃[11]	0.02001	0.02794	0.872	0.00152	0.02642	1.005	0.108
	0.02501	0.02791	0.861	0.00154	0.02637	1.007	0.121
	0.03333	0.02794	0.849	0.00156	0.02638	1.007	0.091
	0.05012	0.02790	0.816	0.00163	0.02627	1.011	0.095
	0.1003	0.02780	0.789	0.00168	0.02612	1.016	0.069
	0.2009	0.02749	0.754	0.00176	0.02573	1.032	0.068
	0.3360	0.02707	0.737	0.00180	0.02527	1.051	0.065
	0.5072	0.02646	0.734	0.00181	0.02465	1.077	0.064
	0.7656	0.02556	0.749	0.00177	0.02379	1.116	0.062
	1.030	0.02458	0.773	0.00172	0.02286	1.161	0.063
KCl[11]	0.02004	0.02791	0.872	0.00152	0.02639	1.006	0.130
	0.02540	0.02791	0.861	0.00154	0.02637	1.007	0.119
	0.03349	0.02773	0.849	0.00156	0.02617	1.015	0.193
	0.05008	0.02767	0.816	0.00163	0.02604	1.020	0.172
	0.1001	0.02727	0.784	0.00169	0.02558	1.038	0.162
	0.2008	0.02654	0.741	0.00179	0.02475	1.073	0.152
	0.3339	0.02566	0.723	0.00183	0.02383	1.114	0.140
	0.5074	0.02466	0.716	0.00185	0.02281	1.164	0.130
KNO₃[11]	0.02002	0.02795	0.872	0.00152	0.02643	1.004	0.087
	0.02501	0.02796	0.861	0.00154	0.02642	1.005	0.087
	0.03336	0.02798	0.849	0.00156	0.02642	1.005	0.087
	0.05011	0.02800	0.816	0.00163	0.02637	1.007	0.061
	0.1003	0.02803	0.784	0.00169	0.02634	1.008	0.035
	0.2011	0.02802	0.741	0.00179	0.02623	1.012	0.026
	0.3340	0.02787	0.723	0.00183	0.02604	1.020	0.026
	0.5182	0.02768	0.716	0.00185	0.02583	1.028	0.023
	0.7812	0.02731	0.716	0.00185	0.02546	1.043	0.023
	1.030	0.02694	0.723	0.00183	0.02511	1.057	0.023

TABLE 1—*Continued*

SALT	μ	m	$\gamma_\pm$	m_+	m_u	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$
BaCl$_2$[5]	0.03441	0.02789	0.850	0.00156	0.02633	1.008	0.101
	0.07629	0.02775	0.795	0.00167	0.02608	1.018	0.102
	0.1425	0.02744	0.763	0.00174	0.02570	1.033	0.099
	0.2998	0.02668	0.720	0.00184	0.02484	1.069	0.097
	0.6048	0.02503	0.686	0.00193	0.02310	1.149	0.100
Ba(NO$_3$)$_2$[5]	0.03162	0.02817	0.853	0.00155	0.02662	0.997	−0.036
	0.07635	0.02823	0.795	0.00167	0.02659	0.998	−0.009
	0.1202	0.02821	0.761	0.00174	0.02647	1.003	0.011
	0.1827	0.02803	0.742	0.00179	0.02624	1.012	0.019
	0.3015	0.02785	0.718	0.00185	0.02600	1.021	0.030
	0.5130	0.02753	0.694	0.00191	0.02562	1.036	0.030
	0.6933	0.02712	0.682	0.00194	0.02518	1.054	0.033

* The citations in the tables are to the corresponding papers.

on gases and non-electrolytes was found to be small. Moreover, the contribution to the ionic strength of the weak acid is not large, and the ionic strength of the added salt is directly obtained.

The quotients in the last columns of the tables are plotted in figures 1 to 4 against the square root of the ionic strength of the added salt, again, for convenience, using this quantity rather than the square root of the ionic strength of the solution. These plots will be discussed later.

THE ACTIVITY COEFFICIENT OF THE UNDISSOCIATED MOLECULES OF ACETIC ACID IN THE PRESENCE OF SALTS FROM MEASUREMENTS OF DISTRIBUTION AND OF THE COMPOSITION OF A DISTILLATE

It seemed desirable in view of the importance of the subject to obtain evidence as to the effect of salts on the undissociated molecules of weak acids by an independent method.

Sugden (12) has measured the distribution of acetic acid between amyl alcohol and water.

If R° is the quotient of the concentration in amyl alcohol by the molality in water, and R is the corresponding ratio when the acid is dissolved, not in pure water, but in a salt solution, we shall set

$$\gamma_u = R/R^\circ \qquad (1)$$

In so doing we neglect the small amount of ionization but as Sugden does not give his acid concentrations but only the distribution ratios, it is impossible to correct for dissociation.

TABLE 2

Activity coefficient of the undissociated molecules of ortho-toluylic acid in salt solutions at 25°

$K_{298.1} = 1.30 \times 10^{-4}$, Solubility in water = 0.008783, $m_+ \gamma_\pm = 0.001003$, $m_u{}^\circ = 0.007745$

SALT	μ	m	$\gamma_\pm$	m_+	m_u	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$
NaCl[5]	0.01200	0.008825	0.892	0.001124	0.007701	1.006	0.217
	0.02500	0.008756	0.863	0.001162	0.007594	1.020	0.344
	0.05000	0.008755	0.829	0.001210	0.007545	1.027	0.231
	0.1000	0.008607	0.789	0.001271	0.007336	1.056	0.237
	0.1600	0.008433	0.765	0.001311	0.007122	1.088	0.229
	0.2500	0.008172	0.746	0.001345	0.006827	1.135	0.220
NaNO$_3$[5]	0.01200	0.008872	0.892	0.001124	0.007748	0.999	−0.015
	0.02500	0.008874	0.863	0.001162	0.007712	1.004	0.069
	0.05000	0.008902	0.829	0.001210	0.007692	1.007	0.061
	0.1000	0.008860	0.789	0.001271	0.007589	1.021	0.090
	0.1600	0.008779	0.765	0.001311	0.007468	1.037	0.099
	0.2500	0.008710	0.746	0.001345	0.007365	1.052	0.088
BaCl$_2$[5]	0.03600	0.008806	0.842	0.001191	0.007615	1.017	0.203
	0.07500	0.008806	0.794	0.001263	0.007543	1.027	0.154
	0.1494	0.008730	0.752	0.001334	0.007396	1.047	0.134
	0.3000	0.008451	0.718	0.001397	0.007054	1.098	0.138
	0.4800	0.008140	0.697	0.001439	0.006701	1.156	0.131
	0.7500	0.007681	0.680	0.001475	0.006206	1.248	0.128
Ba(NO$_3$)$_2$[5]	0.03600	0.008926	0.842	0.001186	0.007740	1.001	0.012
	0.07500	0.008976	0.794	0.001241	0.007735	1.001	0.006
	0.1500	0.008987	0.752	0.001325	0.007662	1.011	0.032
	0.3000	0.008919	0.718	0.001382	0.007537	1.028	0.040
	0.4800	0.008791	0.697	0.001401	0.007390	1.048	0.042
	0.7500	0.008597	0.630	0.001403	0.007194	1.077	0.043

From distillation experiments, McBain and Kam (3) give a quantity which is proportional to the partial pressure of the acetic acid over its solution divided by the concentration of undissociated molecules in the solution. From these the values of the

activity coefficient of the undissociated molecules are obtained directly. The values of γ_u and of $(\log \gamma_u)/\mu(\text{salt})$ for acetic acid in various salt solutions are given in table 5. The first column

TABLE 3

Activity coefficient of undissociated molecules of salicylic acid in salt solutions at 25°

$K_{298\cdot1} = 1.06 \times 10^{-3}$, Solubility in water $= 0.01602$, $m_+ \gamma_\pm = 0.003591$, $m_u^\circ = 0.01218$

SALT	μ	m	$\gamma_\pm$	m_+	m_u	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$
NaCl[11]	0.02003	0.01616	0.867	0.00414	0.01202	1.013	0.280
	0.02502	0.01621	0.858	0.00418	0.01203	1.012	0.207
	0.03331	0.01629	0.845	0.00425	0.01204	1.012	0.156
	0.05000	0.01621	0.827	0.00434	0.01187	1.026	0.223
	0.0999	0.01614	0.789	0.00455	0.01159	1.051	0.216
	0.2499	0.01570	0.746	0.00481	0.01089	1.118	0.194
	0.4927	0.01463	0.733	0.00490	0.00973	1.252	0.198
	0.7325	0.01363	0.745	0.00482	0.00881	1.383	0.192
	0.9714	0.01270	0.767	0.00468	0.00802	1.519	0.187
KCl[11]	0.02002	0.01622	0.867	0.00414	0.01208	1.008	0.173
	0.02503	0.01626	0.858	0.00418	0.01208	1.008	0.138
	0.03331	0.01631	0.845	0.00425	0.01206	1.010	0.130
	0.05000	0.01631	0.827	0.00434	0.01197	1.018	0.155
	0.1000	0.01634	0.784	0.00458	0.01176	1.036	0.154
	0.2492	0.01608	0.732	0.00491	0.01117	1.090	0.150
	0.4915	0.01474	0.717	0.00500	0.00974	1.251	0.198
	0.7495	0.01470	0.716	0.00501	0.00969	1.257	0.133
	1.004	0.01399	0.724	0.00496	0.00903	1.348	0.129
KNO$_3$[11]	0.02003	0.01630	0.867	0.00414	0.01216	1.002	0.043
	0.02504	0.01635	0.858	0.00418	0.01217	1.000	0.000
	0.03324	0.01643	0.845	0.00424	0.01219	0.999	−0.011
	0.05000	0.01654	0.827	0.00434	0.01220	0.998	−0.014
	0.1004	0.01674	0.784	0.00458	0.01216	1.002	0.009
	0.2529	0.01710	0.728	0.00493	0.01217	1.000	0.000
	0.5052	0.01740	0.717	0.00501	0.01239	0.983	−0.015
	0.7528	0.01755	0.716	0.00501	0.01254	0.971	−0.017
	1.004	0.01766	0.724	0.00495	0.01271	0.958	−0.018

gives the added salt, the second the ionic strength of the added salt, the third the activity coefficient of the undissociated acid and the last the quotient of this quantity by the ionic strength μ.

The results are plotted in the usual manner in figure 5. The solid curves are for Sugden's measurements at 25° and the dotted

TABLE 4

Activity coefficient of undissociated molecules of ortho-nitrobenzoic acid in salt solutions at 25°

$K_{298\text{-}1} = 6.12 \times 10^{-3}$, solubility in water $= 0.04415$, $m_+ \ \gamma_\pm = 0.01334$, $m_u{}^\circ = 0.02908$

SALT	μ	m	$\gamma_\pm$	m_+	m_u	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$
NaCl[11]	0.02005	0.04453	0.849	0.01571	0.02882	1.009	0.194
	0.02502	0.04466	0.842	0.01584	0.02882	1.009	0.156
	0.03336	0.04470	0.832	0.01603	0.02867	1.014	0.181
	0.05003	0.04482	0.816	0.01635	0.02847	1.021	0.180
	0.1002	0.04493	0.781	0.01708	0.02785	1.044	0.187
	0.2513	0.04396	0.746	0.01788	0.02608	1.115	0.188
	0.4967	0.04155	0.733	0.01820	0.02335	1.245	0.192
	0.7518	0.03926	0.747	0.01786	0.02140	1.359	0.177
	0.9989	0.03637	0.770	0.01732	0.01905	1.527	0.184
NaNO₃[5]	0.02001	0.04485	0.849	0.01571	0.02914	0.998	−0.046
	0.02510	0.04505	0.842	0.01584	0.02921	0.996	−0.078
	0.03338	0.04529	0.832	0.01603	0.02926	0.994	−0.085
	0.05015	0.04560	0.812	0.01635	0.02925	0.994	−0.054
	0.09986	0.04610	0.781	0.01708	0.02902	1.002	0.009
	0.2509	0.04676	0.746	0.01788	0.02888	1.007	0.012
BaCl₂[5]	0.03726	0.04544	0.820	0.01627	0.02917	0.997	−0.036
	0.07317	0.04564	0.783	0.01704	0.02860	1.017	0.100
	0.1499	0.04586	0.750	0.01779	0.02807	1.036	0.103
	0.2963	0.04577	0.719	0.01855	0.02722	1.068	0.096
	0.5040	0.04433	0.694	0.01922	0.02511	1.158	0.126
	0.7383	0.04344	0.680	0.01962	0.02382	1.221	0.118
Ba(NO₃)₂[5]	0.03432	0.04555	0.825	0.01617	0.02938	0.990	−0.140
	0.08943	0.04637	0.770	0.01732	0.02905	1.001	0.005
	0.1494	0.04668	0.750	0.01779	0.02889	1.007	0.020
	0.2973	0.04747	0.719	0.01855	0.02892	1.006	0.009
	0.5109	0.04838	0.696	0.01917	0.02921	0.996	−0.004
	0.7188	0.04849	0.681	0.01959	0.02890	1.006	0.004

curves for McBain and Kam's measurements at 100°. The curves are in general in agreement with those of the non-electrolytes and weak acids previously studied. The salting-out effect

TABLE 5

Activity coefficient of acetic acid in aqueous salt solutions

SALT	μ	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$	SALT	μ	γ_u	$\dfrac{\log \gamma_u}{\mu(\text{salt})}$
MgSO$_4$,[12] 25°	0.4	1.013	0.014	KCl,[3] 100°	0.2520	1.047	0.079
	1.0	1.041	0.018		0.4740	1.066	0.059
	2.0	1.094	0.020		0.7154	1.101	0.058
	4.0	1.229	0.022		0.9498	1.144	0.061
	8.0	1.580	0.025		1.212	1.165	0.055
					2.072	1.240	0.045
LiCl,[12] 25°	0.100	1.017	0.073				
	0.251	1.046	0.078				
	0.505	1.046	0.039	KBr,[12] 25°	0.1004	1.003	0.013
	1.019	1.207	0.080		0.2525	1.008	0.014
	1.560	1.326	0.079		0.510	1.016	0.014
					1.041	1.035	0.014
NaCl,[12] 25°	0.1001	1.013	0.056		2.164	1.071	0.014
	0.251	1.038	0.065				
	0.5045	1.075	0.062	K$_2$SO$_4$,[12] 25°	0.3768	1.004	0.005
	1.018	1.162	0.064		0.7575	1.017	0.010
	2.072	1.348	0.063		1.532	1.059	0.016
					0.1256	0.994	−0.021
NaCl,[3] 100°	0.1202	1.022	0.079	KNO$_3$,[12] 25°	0.2525	0.988	−0.021
	0.2108	1.050	0.101		0.5105	0.975	−0.021
	0.2915	1.058	0.084		1.043	0.957	−0.018
	0.4129	1.063	0.064				
	0.6689	1.053	0.033		0.3048	1.003	0.004
	1.060	1.199	0.074	KSCN,[3] 100°	0.5758	1.033	0.024
	1.415	1.286	0.077		1.809	1.002	0.001
	2.394	1.479	0.071		2.770	1.063	0.010
	0.1050	0.992	−0.033		0.1027	1.005	0.021
	0.1953	0.997	−0.007		0.1027	1.007	0.030
	0.3312	0.992	−0.011		0.1480	1.008	0.024
	0.5286	0.999	−0.001		0.1480	1.009	0.026
Na$_2$SO$_4$,[3] 100°	0.9270	1.023	0.011		0.2110	1.009	0.018
	1.688	1.076	0.019		0.2110	1.010	0.020
	2.161	1.133	0.025		0.2670	1.035	0.056
	2.384	1.168	0.028		0.2670	1.014	0.021
				Sugar,[3] 100°	0.3637	1.017	0.020
	0.1406	1.005	0.016		0.3637	1.018	0.021
NaC$_2$H$_3$O$_2$,[3] 100°	0.5105	0.989	−0.009		0.5685	1.038	0.029
	1.043	0.968	−0.014		0.5685	1.035	0.026
	1.951	0.938	−0.014		1.274	1.104	0.034
					1.274	1.108	0.035
	0.1003	1.006	0.026		3.096	1.349	0.042
	0.2520	1.015	0.026		3.096	1.363	0.043
KCl,[12] 25°	0.5075	1.032	0.027				
	1.031	1.064	0.026				
	2.124	1.145	0.028				

appears to be lower at the lower temperatures, while in the previous cases studied the effect was lower at higher temperatures. The result is inconclusive, however, as the measurements at the two temperatures were by different authors by different methods. The result by McBain and Kam (3) using sodium sulfate, is not in agreement with previous results. Sugden's (12) results show no abnormal effect with sulfates, such as is found with stronger acids, which will be discussed later.
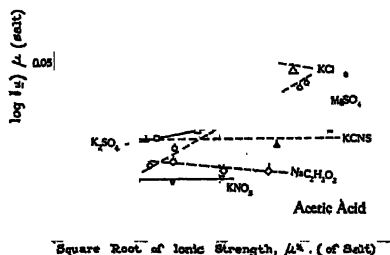

FIG. 5. ACTIVITY COEFFICIENT OF UNDISSOCIATED ACETIC ACID

DISTRIBUTION OF MONOCHLORO- AND DICHLOROACETIC ACIDS BETWEEN NORMAL DI-BUTYL ETHER AND AQUEOUS SALT SOLUTIONS

We will now consider experimentally the activity coefficient of two moderately strong acids. The distribution ratio was used to study the activity coefficient of the undissociated part of monochloro- and dichloroacetic acids. The former is weaker than o-nitrobenzoic acid, the strongest acid used in the previous studies, but the latter is much stronger. Thus, measurements with these acids, by an accurate method, make possible a more severe test of our assumptions of the mechanism of the dissociation of weak acids than has heretofore been possible.

Normal di-butyl ether was found to be the reference solvent best suited to our purpose.

Normal di-butyl ether obtained from the Eastman Kodak Company was treated successively with dilute sulfuric acid, sodium carbonate, and metallic sodium and then subjected to a series of fractional distillations. The fraction which distilled between 141.5 and 142.8° was used.

The monochloroacetic acid was a student preparation. Its melting point was 63°, and the equivalent weight corresponded to the formula.

The dichloroacetic acid was obtained from the Eastman Kodak Company. Considerable difficulty was experienced in the final purification. After some ten fractional crystallizations its melting point was 12.5°, and the equivalent weight was 0.7 per cent too low. A Carius halogen determination indicated that the impurity was probably monochloroacetic acid, no inorganic acid being present. Fractional distillation failed to improve the acid but an extended series of fractional crystallizations gave 200 cc. of acid of M.P. 13.00° and an equivalent weight 0.2 per cent low, which was used in the experiments.

CALCULATION OF THE ACTIVITY COEFFICIENT OF THE UNDISSOCIATED PART OF A WEAK ACID FROM DISTRIBUTION EXPERIMENTS

From distribution data the calculation of the activity coefficient of the undissociated molecules is slightly more involved than in the case of solubility determinations in which the activity of the solid phase is constant. We assume that at a given mol fraction,[1] N, in the non-aqueous layer, the activity is fixed. Then $m_\pm \gamma_\pm^2$ is fixed and as $m_+ = m_-$ we have $m_+ \gamma_\pm$ also fixed. Calculating $m_+ \gamma_\pm$ and $m_u°$ for various total molalities of the weak acid in pure aqueous solutions, we plot them against the values

---

[1] The stoichiometric mol fraction of the acid will be designated as N in the non-aqueous layer, the molality of the hydrogen ion as m_+, of the acid ion as m_-, and of the undissociated acid as m_u in the presence of salt and $m_u°$ in pure acid solutions of total stiochiometric molality, m. The activity coefficient of the undissociated part will be called γ_u, and other symbols have the same meaning as those used by Lewis and Randall (Thermodynamics and the Free Energy of Chemical Substances, McGraw-Hill Book Co., New York, 1923).

of N found to be in equilibrium with solutions of these strengths. In the presence of salts we then determine total acid molality, m, in the aqueous layer, and the mol fraction of acid in the non-aqueous layer.

We may assume that the mean activity coefficient of the dissociated part of the acid is the same as that of hydrochloric acid at the same ionic strength. This assumption is not fully justified, for undoubtedly a part of the abnormality of the activity coefficient of hydrochloric acid arises from the "hydration" of the ions. But in moderately concentrated solutions all strong acids behave more or less like hydrochloric acid. The activity coefficients of sodium and potassium acetates were shown by Randall, McBain and White (6) to approximate closely or even exceed those of the corresponding chlorides. The effect of an error in this assumption is greater in the case of dichloroacetic acid, since the proportion of the ions is greater.

Reading $m_+\gamma_\pm$ from the plot of $m_+\gamma_\pm$ vs. N, and dividing by $\gamma_\pm$ of hydrochloric acid we obtain m_+. The molality of the undissociated molecules is $m_u = m - m_+$. From the plot of m_u° against N we then obtain m_u°, the molality of undissociated molecules in equilibrium with the given value of N in pure aqueous solutions, whence

$$\gamma_u = m_u^\circ/m_u \qquad (2)$$

The method of calculation here used makes no assumption of a constant proportionality of the activity of the acid to its mol fraction in the butyl ether phase. The only assumption made is that a given mol fraction in the ether phase represents the same activity of the acid in either pure water or a salt solution, which is in equilibrium.

THE DISTRIBUTION OF MONOCHLORO- AND DICHLORO-ACETIC ACIDS BETWEEN NORMAL DI-BUTYL ETHER AND WATER AT 25°

Our assumption that the activity is determined by the mol fraction N in the non-aqueous layer involves the assumption that neither water nor salts enter the butyl ether phase, so as to change the activity of the acid. The absence of salts was in all cases

proved by flame tests and tests for the anions. The solubility of dibutyl ether in pure water or in water containing the acids in the concentrations used in the experiments was in all cases less than 0.1 per cent.

The solubility of water in an ether solution in equilibrium with M monochloro-acetic acid is 0.25 per cent. The solubility of

TABLE 6

Distribution of monochloroacetic acid between water and normal di-butyl ether at 25°

$D \frac{25°}{4°}$	m	N	$D \frac{25°}{4°}$	m	N
0.994	0.07867	0.003688	0.996	0.1410	0.006808
0.995	0.08840	0.004167	0.999	0.2152	0.01057
0.995	0.08924	0.004146	0.999	0.2206	0.01083
0.995	0.08979	0.004211	1.009	0.5202	0.02580
0.995	0.09970	0.004724	1.010	0.5483	0.02689
0.995	0.1025	0.004896	1.010	0.5502	0.02729
0.995	0.1030	0.004899	1.010	0.5590	0.02771
0.995	0.1062	0.005056	1.023	1.047	0.05080
0.996	0.1356	0.006542	1.025	1.067	0.05155
0.997	0.1373	0.006626	1.025	1.096	0.05292

TABLE 7

Distribution of dichloroacetic acid between water and normal di-butyl ether at 25°

$D \frac{25°}{4°}$	m	N	$D \frac{25°}{4°}$	m	N
0.994	0.04222	0.007855	1.000	0.1400	0.04129
0.994	0.04263	0.007946	1.001	0.1601	0.04861
0.997	0.07682	0.01862	1.001	0.1654	0.05033
0.997	0.08928	0.02283	1.003	0.2055	0.06574
0.998	0.1068	0.02885	1.004	0.2213	0.07182
0.999	0.1225	0.03471			

water in dibutyl ether containing dichloroacetic acid of about the concentration with which we worked is less than 0.3 per cent. The mol fraction of anhydrous monochloroacetic acid in dry normal di-butyl ether at 0° was 0.270, and with 0.25 per cent water present in the ether the mol fraction (neglecting the water in the ether phase) was 0.275. These values are identical within the limits of error of these crude experiments.

Some of the equilibria were obtained by agitating the layers gently by hand, the greater part were mechanically rotated, and gave the same result. Samples were driven by air pressure through glass tubes, containing a plug of cotton wool into the weighing bottles. The first 30 cc. were rejected in all cases. The distribution experiments with pure acid were repeated at the end of the series. From the fact that the distribution ratios were

TABLE 8

Dissociation constant of monochloroacetic acid[13] at 25°

$K_{298.1} = 0.00139$

c	Λ	Λ/Λ°	$\Lambda_{m^+}+\Lambda_{m^-}$	α	$c\alpha^2/(1-\alpha)$	$\gamma_{\pm}^2$	$\alpha^2 c\gamma_{\pm}^2/(1-\alpha)$
0.000976	265.6	0.688	384.6	0.6906	0.001505	0.943	0.00142
0.001953	219.1	0.569	383.7	0.5710	0.001484	0.928	0.00138
0.003906	174.8	0.453	382.5	0.4566	0.001499	0.910	0.00136
0.007813	136.1	0.353	381.2	0.3570	0.001549	0.891	0.00138
0.01563	103.2	0.267	379.6	0.2719	0.001587	0.871	0.00138
0.03125	77.2	0.200	377.6	0.2044	0.001641	0.847	0.00139
0.0625	56.6	0.147	375.2	0.1509	0.001676	0.821	0.00138

TABLE 9

Dissociation constant of dichloroacetic acid[14] at 25°

c	Λ_o	Λ/Λ°	$\Lambda_{m^+}+\Lambda_{m^-}$	α	$c\alpha^2/(1-\alpha)$	$\gamma_{\pm}^2$	$\alpha^2 c\gamma_{\pm}^2/(1-\alpha)$
0	385.6						
0.003906	359.2	0.932	376.7	0.9535	0.07638	0.876	0.06691
0.007813	338.7	0.887	374.2	0.9051	0.06744	0.838	0.05651
0.01563	309.7	0.803	371.1	0.8346	0.06581	0.795	0.05232
0.03125	273.1	0.708	367.3	0.7435	0.06735	0.749	0.05045
0.06250	231.6	0.601	362.9	0.6382	0.07359	0.706	0.05195
0.1250	190.2	0.493	358.3	0.5309	0.07511	0.662	0.04972

unchanged we may conclude that the butyl ether had not been altered.

The results of the distribution measurements with the pure acids are given in tables 6 and 7. The first column gives the density of the aqueous phase, the second the molality (mols per 1000 grams in vac.) and the third the mol fraction of the acid in the normal di-butyl ether phase.

DISSOCIATION CONSTANT AND CONCENTRATION OF THE
UNDISSOCIATED PART OF THE ACIDS

Before presenting the results of the distribution measurements in aqueous salt solutions we will explain the calculation of the dissociation constants of the acids from conductivity data, and the method of obtaining the molality of the undissociated part of the acids. The calculations of the dissociation constants by the method proposed by Sherrill and Noyes (9) are given in table 8 (Ostwald (13)) table 9, (Kendall (14)), and table 10 (Schreiner (15)). The first columns give the concentration (mols per liter), and the second the equivalent conductance, Λ, in reciprocal international

TABLE 10

Dissociation constant of dichloroacetic acid[15] at 18°

$$K_{291.1} = 0.0583; \quad K_{293.1} = 0.0553$$

c	Λ	$\Lambda/\Lambda°$	$\Lambda_m{}^+ + \Lambda_m{}^-$	α	$c\alpha^2/(1-\alpha)$	$\gamma_\pm^2$	$c\alpha^2\gamma_\pm^2/(1-\alpha)$
0.0	344.8						
0.001	336.7	0.977	342.0	0.9845	0.06253	0.931	0.0582
0.002	332.5	0.965	341.0	0.9751	0.07637	0.906	0.0692
0.005	316.0	0.917	338.9	0.9324	0.06430	0.864	0.0556
0.01	298.3	0.866	336.7	0.8860	0.06886	0.823	0.0567
0.02	274.8	0.797	333.8	0.8162	0.07249	0.778	0.0564
0.05	231.6	0.672	329.0	0.7040	0.08400	0.715	0.0601
0.1	195.2	0.566	324.6	0.6014	0.09074	0.671	0.0609
0.2	157.6	0.457	320.3	0.4920	0.09530	0.632	0.0602

ohms. The third columns give $\Lambda/\Lambda°$ the approximate degree of dissociation. From the first and third columns we obtain the ionic strength of the solution. The fourth column headed $\Lambda_m{}^+ + \Lambda_m{}^-$ gives the sum of the conductances of the positive and negative ions in a solution of the ionic strength equal to the one being measured. The quotient of the corresponding figure of the second column by this quantity gives the true degree of dissociation α (column five). The stoichiometric dissociation constant (column six) is equal to $\alpha^2 c/(1-\alpha)$, and to obtain the thermodynamic or true dissociation constant (column eight) we multiply by the square of the activity coefficient of hydrochloric acid (16) (column seven) in a solution of the same ionic strength.

Schreiner (15) obtained $K_{291}' = 0.0583$ for dichloroacetic acid, which with the heat of dissociation given by Steinwehr (17) gives $K_{298} = 0.0553$ which is in good agreement with the less concordant measurements of Kendall (14). We shall use the value calculated from Steinwehr's measurements.

To obtain the total molality of each acid in equilibrium with various concentrations of its ions we have calculated tables 11 and

TABLE 11

Calculation of undissociated part of monochloroacetic acid in equilibrium with various concentrations of ions

$K_{298.1} = 0.00139$

m_+	$\gamma_\pm$	$(m_+ \gamma_\pm)^2$	$m_u°$	m	x	$m_+\gamma_\pm$
0.0100	0.903	0.00008154	0.05866	0.06866	0.00320	0.00903
0.0125	0.893	0.0001245	0.08957	0.1021	0.00486	0.01116
0.0150	0.886	0.0001766	0.1271	0.1421	0.00684	0.01329
0.0175	0.879	0.0002365	0.1701	0.1876	0.00912	0.01538
0.0200	0.873	0.0003049	0.2194	0.2394	0.01174	0.01746

TABLE 12

Calculation of undissociated part of dichloroacetic acid in equilibrium with various concentrations of its ions

$K_{298.1} = 0.0553$

m_+	$\gamma_\pm$	$m_+ m_- \gamma_\pm^2$	$m_u°$	m	x	$m_+\gamma_\pm$
0.04	0.840	0.001129	0.02042	0.06042	0.0134	0.03360
0.05	0.828	0.001714	0.03099	0.08099	0.0200	0.04140
0.06	0.819	0.002415	0.04367	0.1037	0.0278	0.04914
0.07	0.811	0.003223	0.05828	0.1283	0.0368	0.05677
0.08	0.804	0.004137	0.07481	0.1548	0.0466	0.06432
0.09	0.799	0.005170	0.09350	0.1835	0.0574	0.07190
0.10	0.795	0.006320	0.1143	0.2143	0.0691	0.0795

12. The first column gives the molality of hydrogen ion, which is equal to the molality of acid ion, in the solution which we are considering. The second column gives the activity coefficient in a solution of this ionic strength. The third column gives $m_+ m_- \gamma_+ \gamma_- = K m_u \gamma_u$. Dividing by the dissociation constant we obtain the activity of the undissociated molecules.

We have found the salting out effect of acids upon non-elec-

trolytes to be very small. As the effect of salts on non-electrolytes and weak acids was about the same, we, therefore, neglected the salting out effect of the ions of the acid itself upon the activity coefficient of the undissociated part. Any error in this assumption will be partially cancelled out because the mean molality of the acid and hydrogen ions is approximately constant. Therefore, if we take the activity coefficient of the undissociated molecules in pure solutions of the acids as equal to unity, we set $m_+ m_-$ $\gamma_\pm^2/K = m_u°$, (fourth column). The total molality which is the sum of undissociated and dissociated molecules is given in the column headed m. The column N shows the mol fraction of acid in butyl ether found to be in equilibrium with a pure aqueous solution of acid of the given molality, m. The column of $m_+\gamma_\pm$ is also given. Plots are then made of $m_+\gamma_\pm$ against N, and of $m_u°$ against N.

THE DISTRIBUTION OF MONOCHLORO- AND DICHLOROACETIC ACIDS BETWEEN AQUEOUS SALT SOLUTIONS AND NORMAL DIBUTYL ETHER AT 25°

The results of the distribution measurements with the acids in aqueous solutions without a common chloroacetate ion are given in tables 13 and 15. The first columns give the densities, the second the ionic strengths of the salts, the third the molalities of the acids and the fourth their mol fractions in the butyl ether. The fifth columns, $\gamma_\pm$, are the activity coefficients of hydrochloric acid in a mixture of hydrochloric acid and barium chloride of the ionic strength in the second columns and the same fraction of acid as determined by Randall and Breckenridge (8), or the activity coefficient of hydrochloric acid in a mixture with sodium or potassium chloride as determined by Harned and Åkerlöf (18) recalculated to give 0.795 at 0.1 M. The activity coefficient in mixtures with potassium bromide, potassium nitrate, sodium monochloroacetate and sodium dichloroacetate was arbitarily assumed to be the same as that with the chlorides. The mol fraction in the ether phase fixes the activity $m_+m_-\gamma_\pm^2$ and therefore since $m_+ = m_-$ we know $m_+\gamma_\pm$ (columns six). Reading $m_+\gamma_-$ from the plot of $m_+\gamma_\pm$ against N and dividing by the value

TABLE 13

Monochloroacetic acid in aqueous salt solutions at 25°

$D \frac{25°}{4°}$	μ SALT	m	N	$\gamma_\pm$	$m_+\gamma_\pm$	m_+	m_u	$m_u{}^\circ$	γ_u	$\dfrac{\log \gamma_u}{\mu \text{ (salt)}}$
In sodium chloride solutions
0.999	0.100	0.1017	0.004846	0.789	0.01114	0.0141	0.0876	0.0895	1.022	0.094
1.003	0.200	0.1017	0.004915	0.755	0.01120	0.0148	0.0869	0.0907	1.044	0.094
1.015	0.500	0.0980	0.004977	0.733	0.01129	0.0154	0.0826	0.0920	1.114	0.094
1.033	1.000	0.0954	0.005338	0.770	0.01170	0.0152	0.0802	0.0985	1.228	0.089
1.070	2.000	0.1018	0.006956	0.896	0.01339	0.0149	0.0869	0.1288	1.482	0.085
1.102	3.000	0.0898	0.007271	1.094	0.01372	0.0125	0.0773	0.1350	1.746	0.081
In barium chloride solutions
1.008	0.200	0.1084	0.005177	0.740	0.01154	0.0156	0.0928	0.0955	1.029	0.062
1.025	0.497	0.0991	0.004829	0.705	0.01115	0.0158	0.0833	0.0893	1.072	0.061
1.054	0.988	0.1013	0.005230	0.690	0.01154	0.0167	0.0846	0.0965	1.141	0.059
1.108	1.950	0.0995	0.005693	0.720	0.01210	0.0168	0.0827	0.1053	1.275	0.054
In potassium chloride solutions
1.018	0.500	0.1058	0.005038	0.717	0.01139	0.0159	0.0899	0.0931	1.035	0.030
1.018	0.500	0.0991	0.004704	0.717	0.01098	0.0153	0.0838	0.0865	1.032	0.027
1.039	1.000	0.0993	0.004780	0.722	0.01106	0.0153	0.0840	0.0877	1.044	0.019
1.039	1.000	0.1022	0.004984	0.722	0.01131	0.0157	0.0865	0.0922	1.066	0.028
In potassium bromide solutions
1.036	0.500	0.1008	0.004680	0.717	0.01094	0.0153	0.0855	0.0865	1.012	0.009
1.075	1.000	0.1036	0.004829	0.722	0.01115	0.0154	0.0882	0.0893	1.012	0.005
In potassium nitrate solutions
1.024	0.500	0.1042	0.004580	0.717	0.01079	0.0150	0.0892	0.0845	0.947	−0.047
1.052	1.000	0.1074	0.004576	0.722	0.01079	0.0149	0.0925	0.0843	0.911	−0.041

TABLE 14

Monochloroacetic acid in sodium monochloroacetate solutions at 25°

$D \frac{25°}{4°}$	μ NaAcCl	m	N	$\gamma_\pm^2$	$m_+m_-\gamma_\pm^2\times10^4$	$m_+m_-\times10^4$	m_+	m_u	$m_u{}^\circ$	γ_u	$\dfrac{\log \gamma_u}{\mu \text{ (salt)}}$
1.009	0.2010	0.1265	0.006699	0.552	1.725	3.125	0.0015	0.1250	0.1246	0.997	−0.006
1.028	0.527	0.0897	0.004729	0.512	1.210	2.363	0.0005	0.0892	0.0875	0.980	−0.017
1.057	1.032	0.0994	0.005054	0.523	1.290	2.467	0.0002	0.0992	0.0934	0.942	−0.025

of $\gamma_\pm$ given in the table we obtain m_+ (columns seven), the molality of hydrogen ion. We subtract m_+ from the total molality m and obtain m_u, the molality of undissociated acid. (Columns eight). But from our $m_u{}^\circ$ vs. N plot we find that a pure aqueous solution

TABLE 15

Dichloroacetic acid in aqueous salt solutions at 25°

$D\frac{25°}{4°}$	μ_{SALT}	m	N	$\gamma_\pm$	$m_+\gamma_\pm$	m_+	m_u	$m_u{}^\circ$	γ_u	$\dfrac{\log \gamma_u}{\mu\,(salt)}$
In sodium chloride solutions										
1.005	0.100	0.1651	0.05013	0.758	0.0669	0.0883	0.0768	0.0807	1.051	0.216
1.009	0.200	0.1703	0.05287	0.741	0.0689	0.0930	0.0773	0.0856	1.107	0.221
1.020	0.500	0.1644	0.05424	0.736	0.0698	0.0948	0.0696	0.0879	1.263	0.203
1.038	1.000	0.1487	0.05526	0.779	0.0705	0.0905	0.0582	0.0898	1.543	0.188
In barium chloride solutions										
1.013	0.200	0.1843	0.05624	0.735	0.0711	0.0968	0.0875	0.0915	1.046	0.097
1.027	0.497	0.1751	0.05327	0.715	0.0692	0.0968	0.0783	0.0863	1.103	0.086
1.031	0.497	0.1699	0.05146	0.715	0.0679	0.0950	0.0749	0.0830	1.108	0.090
1.058	0.988	0.1638	0.05254	0.710	0.0687	0.0968	0.0670	0.0850	1.268	0.104
1.112	1.950	0.1557	0.05764	0.745	0.0721	0.0968	0.0589	0.0939	1.594	0.104
In potassium nitrate solutions										
1.031	0.500	0.1751	0.05101	0.715	0.0675	0.0944	0.0807	0.0824	1.021	0.018
1.059	1.000	0.1782	0.05193	0.726	0.0682	0.0939	0.0841	0.0838	0.996	−0.002

TABLE 16

Dichloroacetic acid in potassium dichloroacetate solutions at 25°

$D\frac{25°}{4°}$	μ KAcCl$_2$	m	N	$\gamma_\pm^3$	$\dfrac{m_+m_-}{\gamma_-}$	$\dfrac{m_+m_-}{}$	m_+	m_u	$m_u{}^\circ$	γ_u	$\dfrac{\log \gamma_u}{\mu\,(salt)}$
1.007	0.1083	0.1201	0.04550	0.599	0.004032	0.006729	0.0440	0.0761	0.0728	0.09566	−0.178
1.018	0.2667	0.0880	0.03815	0.532	0.003352	0.006301	0.0218	0.0662	0.0610	0.09215	−0.133
1.047	0.6273	0.0753	0.03391	0.510	0.002970	0.005824	0.0915	0.0661	0.0535	0.08094	−0.146

of the acid of a molality $m_u{}^\circ$ (columns nine) in undissociated molecules is also in equilibrium with the mol fraction N in butyl ether and therefore has the same activity. Then, $m_u \gamma_u = m_u{}^\circ \gamma_u{}^\circ$. As $\gamma_u{}^\circ$ the activity coefficient of undissociated molecules in a

pure solution of the acid is equal to unity we have $\gamma_u = m_u{}^\circ/m_u$. Columns 10 give the activity coefficient of undissociated acid molecules, γ_u. Columns 11 give the logarithm of the activity coefficient of undissociated acid molecules divided by the ionic strength of the salt solution.

Tables 14 and 16 present measurements made in the presence of a common anion. The calculation is very similar. We ob-
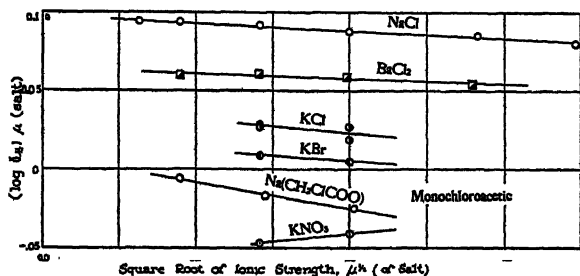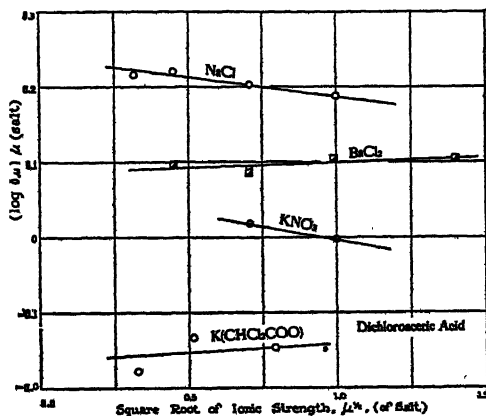

FIG. 6. ACTIVITY COEFFICIENT OF UNDISSOCIATED MONOCHLOROACETIC ACID

FIG. 7. ACTIVITY COEFFICIENT OF UNDISSOCIATED DICHLOROACETIC ACID

serve m and N. From the plot against N we obtain $m_+\gamma_\pm$ in pure acid solution in equilibrium with a butyl ether phase of the same mol fraction. Squaring this, one obtains $(m_+\gamma_\pm)^2$ in pure solution and this equals $m_+m_-\gamma_\pm{}^2$ in any solution of the same activity with or without a common ion. We have given the

molality of acid anion added as the salt. By successive approximations we find m_+ and m_- such that m_- (total) $-m_-$ (added anion) $= m_+$ and m_- (total) $\cdot m_+\gamma_\pm{}^2 = m_+m_-\gamma_\pm{}^2$ obtained above. Then $m_u = m - m_+$ and again $\gamma_u = m_u{}^\circ/m_u$.

The results are plotted in figures 6 and 7 in which the ordinate is the quotient of the logarithm of the activity coefficient of the undissociated molecules by the ionic strength of the salt. For convenience the square root of the ionic strength of the added salt rather than of the solution was used as the abscissae of the plots.

DISCUSSION OF THE ACTIVITY COEFFICIENT OF THE UNDISSOCIATED PART OF MONOCHLORO- AND DICHLOROACETIC ACIDS

The plots of figures 6 and 7 show a striking similarity to similar plots of the gases, non-electrolytes and weak acids studied in the previous papers. The magnitude of the salting out effect is small with acetic acid, is slightly larger with monochloro- and largest with dichloroacetic acid.

The curves are nearly horizontal, and therefore give additional evidence as the validity of the prediction made in the earlier paper that the quotient of the logarithm of the activity coefficient of a non-electrolyte or of the undissociated part of a weak acid by the ionic strength of the added salt is approximately a constant. However, we must remember that certain assumptions regarding the ions were made, namely, that the activity coefficient of the dissociated ions was the same as that of hydrochloric acid at the same ionic strength, and second, that the salting out effect of the dissociated part of the acid itself was negligible. The dissociated part of the acid contributes from 0.4 to 12 per cent in the case of monochloro- and from 5 to 40 per cent in the case of the dichloroacetic acid. The salting out effect of the common chloroacetate ion is negative, for the effect of sodium and potassium ions is predominately positive. We have not, however, attempted to make a correction for this effect.

We have also neglected the activity coefficient of the undissociated part of the acid in the pure aqueous solutions. In the case of acetic acid it is small, but its magnitude in the substituted acids is unknown.

THE ACTIVITY OF WEAK ACIDS IN AQUEOUS SULFATE SOLUTIONS

When sulfates are added to solutions of weak acids, there is a distribution of the hydrogen ion constituent between the weak acid ion and the hydrosulfate ion. We have as yet insufficient data to fully consider such cases. We show in figure 8 the ratio of the molality of several acids in pure water, to the molality in a sulfate solution in which its activity has the same value (19).
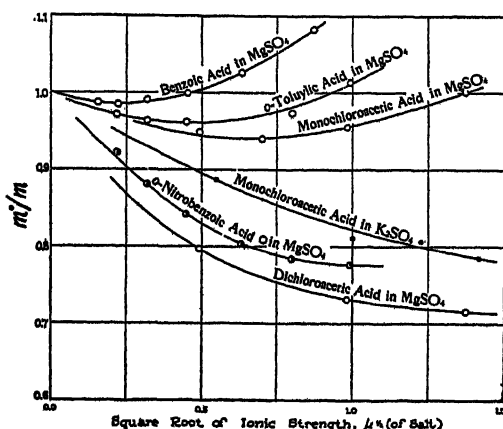

FIG. 8. "ACTIVITY COEFFICIENT" AT CONSTANT ACTIVITY OF WEAK ACIDS IN AQUEOUS SULFATE SOLUTIONS AT 25°C.

GENERAL DISCUSSION

The results of this and the preceding papers are summarized in table 17. In this table we have assembled the data under the head of the added salt, which is given in the first column. The second column gives the non-electrolyte or weak acid, and the third the average value of the quotient of the logarithm of the activity coefficient by the ionic strength of the added salt.

In examining the table we find that the salting-out effect of the salts is about the same upon the undissociated part of the weak acids as upon gases and non-electrolytes. The effect is small for acetic, somewhat larger for monochloroacetic, and largest for dichloroacetic acid. As was pointed out in the article dealing with gases, if the relation of Debye and McAulay is true

TABLE 17

Summary of all calculations on activity coefficients of undissociated weak acids and non-electrolytes

ADDED SALT	SUBSTANCE	$\frac{\log \gamma}{\mu(\text{salt})}$	ADDED SALT	SUBSTANCE	$\frac{\log \gamma}{\mu(\text{salt})}$
HCl	O_2	0.036	AlCl₃	H_2	0.028
	H_2	0.020		C_2H_2	0.044
	N_2O	0.016		b	0.070
	CO_2	0.009			
	H_2S	0.014	Al₂(SO₄)₃	N_2O	0.046
	a	−0.036		C_2H_2	0.043
H₂SO₄	O_2	0.030	Al(NO₃)₃	N_2O	0.045
	H_2	0.018		C_2H_2	0.023
	CO_2	0.017			
	H_2S	0.052	MgCl₂	C_2H_2	0.055
HNO₃	O_2	0.023			
	H_2	0.007	MgSO₄	H_2	0.057
	N_2O	−0.014		N_2O	0.070
	CO_2	−0.022		C_2H_2	0.064
	a	−0.025		b	0.075
ZnSO₄	N_2O	0.061	Mg(NO₃)₂	N_2O	0.035
	C_2H_2	0.061		C_2H_2	0.026
Zn(NO₃)₂	N_2O	0.042	CaCl₂	H_2	0.065
	C_2H_2	0.031		N_2O	0.064
Cu(NO₃)₂	N_2O	0.065		C_2H_2	0.058
MnSO₄	N_2O	0.059	Ca(NO₃)₂	N_2O	0.040
	C_2H_2	0.056		C_2H_2	0.027
FeSO₄	N_2O	0.033	BaCl₂	N_2	0.120
	C_2H_2	0.061		N_2O	0.073
FeCl₃	C_2H_2	0.030		CO_2	0.060
				C_2H_2	0.065
Fe₂(SO₄)₃	N_2O	0.027		b	0.075
	C_2H_2	0.027		c	0.100
CoSO₄	N_2O	0.081		d	0.134
	C_2H_2	0.061		e	0.105
				ClAc	0.056
NiSO₄	N_2O	0.080		Cl₂Ac	0.098
	C_2H_2	0.066	Ba(NO₃)₂	c	0.030
				d	0.040
Cr₂(SO₄)₃	N_2O	0.022		e	0.004
	C_2H_2	0.021	LiOH	NH_3	0.080

TABLE 17—*Continued*

ADDED SALT	SUBSTANCE	$\dfrac{\log \gamma}{\mu(\text{salt})}$	ADDED SALT	SUBSTANCE	$\dfrac{\log \gamma}{\mu(\text{salt})}$
LiCl	H_2	0.066	Na$_2$SO$_4$	N_2O	0.110
	N_2O	0.084		H_2S	0.180
	NH_3	−0.020		C_2H_2	0.100
	Ac	0.075		I_2	0.107
LiBr	NH_3	−0.035		b	0.130
LiI	NH_3	−0.062			
LiNO$_3$	b	−0.012	NaNO$_3$	H_2	0.082
NaOH	O_2	0.188		N_2O	0.072
	H_2	0.140		H_2S	0.034
	NH_3	0.090		C_2H_2	0.050
NaCl	O_2	0.133		I_2	0.045
	H_2	0.094		a	−0.130
	N_2	0.200		b	0.020
	N_2O	0.100		c	0.064
	H_2S	0.060		d	0.089
	NH_3	0.027		f	−0.078
	C_2H_2	0.093	NaH$_2$PO$_4$	I_2	0.140
	a	0.089	NaAc	Ac	−0.014
	b	0.013	NaClAc	ClAc	−0.016
	c	0.191	KOH	O_2	0.175
	d	0.232		H_2	0.120
	e	0.180		N_2O	0.130
	f	0.196		NH_3	0.120
	Ac	0.066			
	ClAc	0.088	KCl	H_2	0.080
	Cl$_2$Ac	0.204		N_2O	0.081
NaClO$_3$	b	0.033		CO_2	0.063
NaClO$_4$	b	−0.020		H_2S	0.054
NaBr	N_2O	0.090		NH_3	0.035
	H_2S	0.016		C_2H_2	0.064
	NH_3	−0.014		a	0.034
	C_2H_2	0.080		b	0.120
				c	0.152
				f	0.140
NaI	NH_3	−0.025		Ac	0.033
	b	−0.030		ClAc	0.026
			KClO$_3$	NH_3	0.064
				b	0.020

TABLE 17—*Concluded*

ADDED SALT	SUBSTANCE	$\frac{\log \gamma}{\mu(\text{salt})}$	ADDED SALT	SUBSTANCE	$\frac{\log \gamma}{\mu(\text{salt})}$
KBr	N_2O	0.068	KCl_2Ac	Cl_2Ac	−0.150
	CO_2	0.050	KSCN	Ac	0.010
	H_2S	0.006	RbCl	N_2O	0.076
	NH_3	0.062		CO_2	0.050
	C_2H_2	0.050	$RbNO_3$	b	−0.050
	b	0.055	CsCl	N_2O	0.058
	Ac	0.014		CO_2	0.042
	ClAc	0.008	$CsNO_3$	b	−0.130
$KBrO_3$	NH_3	0.043	NH_4Cl	N_2O	0.030
KI	N_2O	0.015		CO_2	0.020
	CO_2	0.039		H_2S	0.001
	H_2S	−0.012		C_2H_2	0.023
	NH_3	−0.010	NH_4Br	N_2O	0.024
	b	−0.060		H_2S	−0.018
KIO_3	NH_3	0.023		C_2H_2	0.012
KIO_4	N_2O	0.055	$(NH_4)_2SO_4$	C_2H_2	0.045
K_2SO_4	O_2	0.104	NH_4NO_3	H_2	0.010
	H_2S	0.130		N_2O	0.008
	NH_3	0.055		H_2S	0.021
	C_2H_2	0.082		b	−0.070
	b	0.120	g	H_2	0.000
KNO_3	H_2	0.062		N_2O	−0.017
	N_2O	0.047		CO_2	0.022
	CO_2	0.031	Sugar	H_2	0.050
	H_2S	0.018		CO_2	−0.010
	NH_3	0.033		Ac	0.025
	C_2H_2	0.030	Urea	N_2	0.000
	a	−0.160		N_2O	−0.004
	b	0.000		H_2S	−0.027
	c	0.025			
	f	−0.006			
	Ac	−0.020			
	ClAc	−0.043			
	Cl_2Ac	0.010			
K_2CO_3	H_2	0.090			

(*a*) *o*-nitrobenzaldehyde; (*b*) phenylthiourea; (*c*) benzoic acid; (*d*) *o*-toluylic acid; (*e*) *o*-nitrobenzoic acid; (*f*) salicylic acid; (*g*) chloral hydrate; (Ac) acetic acid; (ClAc) monochloroacetic acid; (Cl_2Ac) dichloroacetic acid.

then the ratio of $(\log \gamma)/\mu$ for one non-electrolyte to $(\log \gamma)/\mu$ for another non-electrolyte within the group should be the same as the ratio for the same non-electrolytes in another group. We note that this is qualitatively the rule for the weak acids as well as for non-electrolytes.

It is difficult to interpret negative values of $(\log \gamma)/\mu$. In general nitrates and organic anions give very small or negative values of $(\log \gamma)/\mu$ and the bromides and iodides in several cases are negative. The iodides have large ionic radii in the Debye and Hückel sense, while the ionic radius of nitrates is thought to be quite small (20).

It must be remembered, however, that to some extent the values of the function which we are considering are dependent on the units chosen for expressing concentrations. For example if we were to use the ratio of the activity of the undissociated substance to its mol fraction, then the activity coefficient would have different values in moderately concentrated solutions. Further generalizations do not seem possible at the present time.

In the case of the added sulfates the quantity $m°/m$ defines a quantity which is analogous to the activity coefficient for solutions of the same activity. The quantity $m°/m$ is lowered by a larger amount the larger the dissociation constant of the acid. This is in agreement with an assumption of a larger amount of hydrosulfate ion formed with the stronger acid. The ordinary salting-out effect, which is pronounced in the case of sulfates on non-electrolytes, is more noticeable with the weaker acids.

GENERAL SUMMARY

The activity coefficient of gases, of solid non-electrolytes, and of the undissociated part of weak acids, has been considered as the measure of the deviation of the properties of these substances from the laws of the perfect solution.

The quotient of the logarithm of the activity coefficient of the gases and of the non-electrolytes by the quotient of the ionic strength in aqueous salt solutions is approximately constant.

This constant varies with the salt or the non-electrolyte being considered.

The value of this constant, which may be called the salting-out effect, may be used as the basis for arranging a salting-out series of the ions. In the series so arranged, the valence of the ion has little effect in determining the order.

The distribution of monochloro- and dichloroacetic acid between aqueous salt solutions and normal di-butyl ether has been determined.

If the simple assumption that the dissociated part of weak acids acts in the presence of added salts in the same way as the ions of hydrochloric acid, then the quotient of the logarithm of the activity coefficient of the undissociated part by the ionic strength of the added salt is also approximately constant. The salting-out order of the ions is also the same as that for non-electrolytes.

When sulfates are added to solutions of weak acids, some hydrosulfate ion is formed. The results of partition and solubility measurements in the presence of sulfates are qualitatively explained by this assumption and effects similar to the above.

A general summary of the effect of individual electrolytes on the several non-electrolytes is given.

REFERENCES

(1) ARRHENIUS: Z. physik. Chem., 31, 197 (1899).
(2) McBAIN AND COLEMAN: J. Chem. Soc., 105, 1517 (1914).
(3) McBAIN AND KAM: Ibid., 115, 1332 (1919).
(4) BRÖNSTED: J. Chem. Soc., 119, 574 (1921).
(5) RÖRDAM: Thesis, Copenhagen, 1925.
(6) RANDALL, McBAIN AND WHITE: J. Am. Chem. Soc., 48, 2517 (1926).
(7) HARNED: J. Am. Chem. Soc., 48, 326 (1926).
(8) RANDALL AND BRECKENRIDGE: J. Am. Chem. Soc., 49, 1435 (1927).
(9) SHERRILL AND NOYES: J. Am. Chem. Soc., 48, 1861 (1926).
(10) RANDALL AND YOUNG. Unpublished.
(11) HOFFMANN AND LANGEBECK: Z. physik. Chem., 51, 385 (1905).
(12) SUGDEN: J. Chem. Soc., 129, 177 (1926).
(13) OSTWALD: Z. physik. Chem., 3, 170 (1889).
(14) KENDALL: Medd. Vetenskapsakad. Nobelinst., 3, No. 2 (1914).
(15) SCHREINER: Z. anorg. allgem. Chem., 122, 201 (1922).
(16) RANDALL AND YOUNG: Unpublished.
(17) STEINWEHR: Z. physik. Chem., 38, 185 (1901).
(18) HARNED AND ÅKERLÖF: Physik. Z., 27, 411 (1927).
(19) RANDALL AND FAILEY: J. Am. Chem. Soc., 49, No. 11 (1927).
(20) BJERRUM: Det. Kgl. Danske Videns. Selskab, Math.-fys. Medd., 7, No. 9 (1926). .

ANTAGONISM OF RADIATIONS IN PHOTOCHEMICAL AND PHOTOGRAPHIC REACTIONS[1]

S. E. SHEPPARD

The Kodak Research Laboratories, Eastman Kodak Company, Rochester, New York[2]

In the course of testing certain cellulose nitrate products in the form of transparent sheets, a spontaneous formation of yellow color was observed. It was found that this coloration could be bleached out by exposure to sunlight, the time varying with the intensity of the color, the strength of the sunlight, and other factors, but being of the order of several hours. When it was attempted to accelerate the bleaching by exposure of the cellulose nitrate sheet to the light from a mercury lamp in quartz, it was found that not the bleaching but on the contrary the yellow coloration was intensified. In fact, the same coloration which required several months to produce in darkness or ordinary diffused light was now produced in one-half to one hour.

On screening off the extreme ultra-violet rays with glass, it was found that the reaction was reversed, bleaching taking place in the near ultra-violet and in the blue-violet. The yellow and green rays from the mercury lamp were found to play little or no part, but the ultra-violet rays absorbed by cellulose nitrate accelerated the coloration. The following spectrum diagram shows roughly the distribution of the radiation antagonism (fig. 1).

The yellow to deep orange coloration was traced to nitration of aromatic bodies, particularly of a phenolic character, present as impurities or intentionally in the sheet. By removing those accidentally present, and putting in pure phenol, a more definitely controllable reaction was obtained, in which it could be shown

319

that the photochemical decomposition of the cellulose nitrate, liberating N_2O_4 and N_2O_3, was followed by nitration of the phenol. Apparently both mono- and di-nitrophenols were formed, and there was in addition some breakdown of the cellulosic residue itself, as well as more complicated reactions of the phenol, which was acting as acceptor for the nitrogen peroxide.

The near ultra-violet and blue-violet rays absorbed by the nitrophenols tended to reverse the reaction, the cellulose residue being partially renitrated, when the shorter ultra-violet rays were screened off.

Another example of antagonistic photochemical actions of radiations which has recently been studied in this laboratory is the formation of hydrogen peroxide in the oxidation of resin. If paper soaked with sodium resinate, or freshly cut resinous

YELLOWING RAYS   BLEACHING RAYS
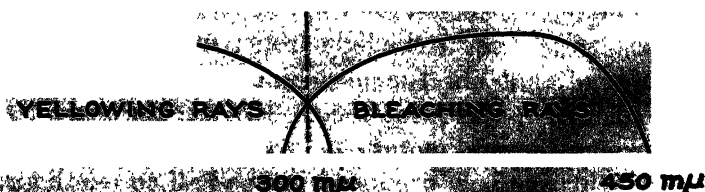
300 mμ   450 mμ

FIG. 1. SPECTRUM DIAGRAM OF RADIATION ANTAGONISM

wood, drying oils, or a number of other autoxidizable substances are placed in contact with a photographic plate, and left for say twelve to twenty-four hours, upon development an image is formed of the material. This effect, known as the Russell effect (1) is due to the formation of hydrogen peroxide, which even in small traces, of the order of 1 in 1,000,000, has a definite fogging effect upon the photographic plate. It has been shown by Sheppard and Wightman (2) that the formation of a latent image by peroxide very closely resembles the photographic action of light. Now the pseudo-photographic action[3] of resins and similar peroxide-forming substances is greatly increased by previous exposure to light. The peroxide-forming reaction is

[a] Term suggested by W. Clark (in a very valuable review of these effects), Science Progress.

photochemically catalyzed. In order to increase the rate of this action, paper soaked with resin was exposed to ultra-violet light from a quartz mercury lamp,[4] the formation of peroxide being followed by its effect upon a photographic plate. The plate was placed in darkness and left for twelve hours in contact with the material, then developed under standard conditions. It was found that when the peroxide forming capacity of the exposed material was small, increased exposure to the quartz ultra-violet produced first an increase of the pseudo-photographic action, then a decrease, or apparent reversal. (Cf. fig. 2.)

FIG. 2. EFFECT OF ULTRA VIOLET LIGHT ON A PAPER IMPREGNATED WITH SODIUM RESINATE

Clear. Increasing exposure, right to left

FIG. 3. EFFECT OF ULTRAVIOLET LIGHT ON A PAPER IMPREGNATED WITH SODIUM RESINATE

Wratten filters 35 + 45. Increasing exposure, right to left

On the other hand, when filters cutting out the ultra-violet were used, the plate showed little or no sign of reversal for the same range of exposure (fig. 3).

By using light filters, it was found that the reversing action was eliminated when ordinary glass, absorbing rays below 400 mμ, was placed before the material. It was established that for rosin the maximum peroxide producing action was between 400 and 300 mμ, while rays below 290 mμ reduced the action.

[4] Run at 180 volts, 8 amp., at a distance of 25 cm.

This reversing action is due to the fact that the short ultra-violet rays decompose hydrogen peroxide (3) so that as with the yellowing of cellulose nitrate, a *radiation antagonism* was established between the longer and shorter ultra-violet rays. In this case, the antagonism is due to the destructive action of another part of the spectrum upon the primary photochemical product. There is no reformation of the original reactant, so far as is known. It appeared from this that radiation antagonism is of much importance in practical photochemistry, and invites some discussion and classification. A paper by G. Rabel (4) gives a number of interesting examples, but is chiefly concerned with a possible relation between polar (antagonistic) actions of parts of the spectrum and luminescence stratification in electric discharge through rarefied gases. In the following the subject is discussed more particularly with reference to photochemical reactions.

The idea of antagonistic action of radiations of different frequency[5] is relatively old in photochemistry. It appears to have been first specifically formulated by Ritter (5) at the beginning of the nineteenth century. It had been observed by Wilson that Bonioni's phosphor was made to shine more brightly by violet than by red rays, and that red rays thrown on a phosphor previously excited by violet rays weakened the luminescence. Ritter remarks on this "But the violet rays belong to the reducing, the red to the oxidizing part of the spectrum." This statement of a "chemical polarity" of the spectrum therefore preceded the enunciation of Grotthus' photochemical principles (6) and might be regarded as the first photochemical generalization.

The quenching action of red and infra-red rays upon the phosphorescence of alkaline earth sulfides was confirmed by Ed. Becquerel (7) who further showed that it was to be definitely distinguished from the reviving action of heat [thermo-phosphorescence].[6] Becquerel made few speculations in regard to

---

[5] Originally specified as rays of different refrangibility.

[6] P. Lenard (*Heidelb.*, *Ber.*, 1917, 5; *ibid.*, 1918) considers that part of this quenching effect is a heat effect, whereby acceleration of the phosphorescent decay is brought about, but part is a true quenching effect of entirely different order.

the numerous facts he discovered in photochemistry. But, what was equivalent, he introduced the terms "rayons excitateurs," "rayons continuateurs," "rayons extincteurs," whereby the concept of antagonism of radiations was given a more specific, and in some ways a more misleading, form. Be-

TABLE 1

REACTION	→ RADIATION	← RADIATION	REMARKS
$2 HCl \rightleftharpoons H_2 + Cl_2$ $- Q$	220 $\mu\mu$	Blue and ultra-violet above 250 $\mu\mu$	The extreme ultra-violet rays act oppositely to the near ultra-violet
$2 HBr \rightleftharpoons H_2 + Br_2$ $- Q$ $H_2 + \rightleftharpoons 2 HBr$ $- Q$	209 $\mu\mu$ 263 $\mu\mu$	Violet and blue 207 $\mu\mu$ 253 $\mu\mu$ 282 $\mu\mu$	The direction of energy storage is reversed
$2 SO_2 + O_2 \rightleftharpoons 2 SO_3$ $- Q$	Ultra-violet short	Ultra-violet long	
$CO + Cl_2 \rightleftharpoons COCl_2$ $- Q$	Ultra-violet short	Ultra-violet long	
$2 H_2 + O_2 \rightleftharpoons 2 H_2O$	Ultra-violet short	Ultra-violet long	Oxygen photo-active component. Plotnikow suggests that there are antagonism bands in the spectrum
$2 NH_3 \rightleftharpoons N_2 + 3H_2$ $- Q$	Below 189 $\mu\mu$	203 $\mu\mu$ 254 $\mu\mu$	Equilibrium not attained in quartz, because it absorbs below 185 $\mu\mu$
$3 O_2 \rightleftharpoons 2 O_3$ $- Q$	Below 200 $\mu\mu$	250 $\mu\mu$	Ozone is also decomposed by heat, and is formed by violet and blue light
$SO_2 + Cl_2 \rightleftharpoons SO_2Cl_2$ $- Q$	Extreme ultra-violet	Violet	

consequently was implied that certain rays in the spectrum possess a specific exciting (or activating) effect on chemical reactions, in general, others a continuating effect, and others again an extinguishing or destroying effect. This was in similar terms to the

division of the spectrum into actinic and thermic rays, in which connection it may be noted that long after Grotthus and Draper had founded the absorption law of photochemistry, J. M. Eder (8) and still more definitely J. Plotnikow (9) have assumed that a substance may have some absorption bands which are "thermic" as distinct from others which are "chemical."

In the development of photochemistry, three main classes of reactions definitely exhibit the *phenomenon* of radiation antagonism. These are (a) real reversible photochemical reactions, (b) phosphorescence, (c) phototropy. The typical features of the real reversible photochemical reactions are illustrated in table 1, taken with some amplifications from Plotnikow (10).

For none of these reactions is the actual mechanism completely known[7] and the above tabulation therefore is only a crude representation. However, a certain general feature may be pointed out. The reaction proceeding with absorption of energy (endo-energetic) is produced by the (relatively) higher frequencies, or short wave-lengths, while the opposite, exo-energetic reaction, releasing energy, is accelerated by the (relatively) longer wave-lengths.

The attempt of Perrin (11) to generalize all chemical reactions under a radiation theory of chemical change has been very stimulating, but has fallen short of full success. In Perrin's general equation

$$A_1 + h\nu_1 \rightleftharpoons A_2 + h\nu_2$$
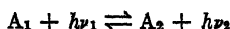
A_1 and A_2 represent different chemical configurations, reactant and resultant, while $h\nu_1$ and $h\nu_2$ represent quanta of monochromatic radiation. Not only in actual photochemical reactions are the exciting radiations extended over a considerable region of wavelengths (12), but any complete equation would have to include kinetic energy terms of *rotation* and *oscillation*—both quantized— and also of translation,—not quantized (13). These terms may refer partly to groups of these,—radicles, ions and molecules, and in addition an important part of the translational, or non-

---

[7] Cf. particularly the recent symposium on photochemical reactions in gases at the Faraday Society, London (October, 1925).

quantized energy may be imparted to entirely non-component atoms or atom groups. Hence it appears that the foregoing equation is an over-simplification when applied to actual chemical reactions.
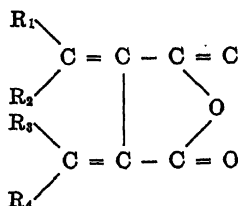
It is evident, however, that this generalized equation is, so to speak, the apotheosis of the Ritter concept of radiation antagonism. Furthermore, it may represent the fundamental or primary element of photochemical change, considered as a virtual change, which is related to real chemical change in the manner suggested by Bohr, Kramers, and Slator (14) for the relation of virtual radiation in general to quantum exchange of energy. In consequence we should expect the nearest approach to fulfillment of the Perrin-McLewis equation in the field, if existing, of *virtual photochemical* reactions. Such fields appear to exist in the case (*a*) of phosphorescence and (*b*) of phototropy. Both these are characteristic of the solid state, phototropy exclusively so. While phosphorescence is a phenomenon of dilute solid solutions,[8] *phototropy* is a property of certain pure substances, and should perhaps be restricted to the reversible light induced color changes of crystallized organic compounds.

It is not proposed to discuss here either the enduring electroluminescence of certain gases (e.g., active nitrogen), or the relation of phosphorescence to fluorescence.

The term phototropy has been used in a wider sense, as referring to any causes of reversible color changes induced by light (15). The existence of phototropy in inorganic substances is not denied, but the extension of the term to phenomena such as the color-adaptation of the photo-halides had better be withdrawn.

Phototropy in this narrow sense appears to have been first observed by Marckwald (16). Later it was more extensively studied by H. Stobbe (17) who found in the *fulgides* a class of colored organic substances showing phototropy in clearly marked fashion. The fulgides have assigned to them the general formula:

[8] But compare W. D. Bancroft's criticism of this term, *Trans. Faraday Soc.*, **19**, 324 (1923).

$$
\begin{array}{c}
R_1 \diagdown \\
\qquad C = C - C = C \\
R_2 \diagup \qquad | \qquad \diagdown \\
R_3 \diagdown \qquad | \qquad \diagup O \\
\qquad C = C - C = O \\
R_4 \diagup
\end{array}
$$

where $R_1R_2R_3R_4$ may be hydrogen, identical, or different alkyl or aryl groups.

Typical phototropic oscillations are:

	Diphenylfulgide	
A form	$\rightleftharpoons$	B form
Yellow green		Blue
Absorbs ca		Absorbs ca
510–436 $\mu\mu$		510–625 $\mu\mu$

	Triphenylfulgide	
A form	$\rightleftharpoons$	B form
Orange		Blue
Absorbs ca		Absorbs ca
550–440 $\mu\mu$		550 $\mu\mu$–infra-red

It is characteristic of phototropic substances to be transformed, by the light they absorb, into a form of more or less complementary hue, which is reciprocally inverted by absorption of light in its own absorption region. In mixed radiation, therefore, an equilibrium, or at least a stationary state, is reached, depending upon the proportions of rays of different frequencies incident.

The nature of phototropic change has not been definitely settled. Attempts to explain it in terms of isomerism, static or dynamic, have not so far been successful, particularly in view of the fact that its limitation to the solid state seems to preclude its being any known form of inter-molecular variation. P. Gallagher (18) in a study of a number of salicylidene amines was not able to trace any direct relation between phototropy and chemical constitution, and concluded that the color changes cannot be attributed to purely chemical changes in the molecule. On the other hand, there is certain evidence difficult to reconcile with entire absence of inter-molecular chemical change. Senier

and Shepheard (19) as also Padoa and his coworkers (20) found constitutional factors involved. Again, Schlunk and Lumpp noted with ozazones that when 3 or 4 of the four hydrogens in the molecule are replaced by acyl groups that phototropy is excluded. This condition would also eliminate or reduce the possibility of "labile" hydrogen, and it is possible that a "coordinated" hydrogen atom (21) may play a part in the phototropic transformation. There are phototropic substances in which the reverse reaction occurs spontaneously in the dark. This reverse reaction has a high temperature coefficient (vide infra) and it seems probable that the reverse change is in fact accelerated by infra-red rays.

The magnitude of the temperature coefficient of the reverse

TABLE 2

(By Padoa and collaborators)

PHOTOTROPIC SUBSTANCE	TEMPERATURE INTERVAL	DIRECT CHANGE T.C.	RE-VERSE CHANGE T.C.	TEMPERATURE INTERVAL
1. Piperyl o-tolyl-ozazone.......	$-10°$ to $-10°$	1.06	2.00	$+10°$ to $-10°$
2. Benzaldehyde phenyl hydrazone........................	$-10°$ to $+10°$	1.00	1.70	$80°$ to $110°$
3. p,p^L-diacetyl-diamino-stilbene-o, o'-disulfonic acid.......	$-10°$ to $+20°$	1.07		
4. salicylidene-β-naphthylamine.	$-10°$ to $0°$	1.40	2.00	

reaction is generally very high for a purely physical reaction (see table 2).

These reactions follow the general rule for photochemical reactions, that the temperature coefficient increases with the wave-length of the active radiation. If the reverse reaction be regarded, in view of its temperature-coefficient, as chemical, it would seem that the direct reaction must be so also. Padoa found that the direct reaction was monomolecular, the reverse bimolecular, suggesting a depolymerization-polymerization (dimerization)

$$A_2 \rightleftharpoons 2 A$$

as the change in these specific cases. The most complete study

of a phototropic change was made by F. Weigert (22) in the
case of β-tetra-chlor-α-keto-naphthalin. He found that with the
exception of the color (absorption) none of the crystallographic
properties underwent any noticeable alteration. The axial ratio
of the rhombic crystals remained practically constant under the
strongest excitation, and the x-ray (Laue) diagram showed no
difference. The absorption spectrum was very different, accord-
ing as the electric vector of the light passing was plane polarized
in the direction of the c-axis (prism edge), or in any direction of
the a–b-plane. For the unexcited crystal, the first direction
gave two absorption bands at 295 and $375\mu\mu$, the second, a general
absorption beginning at $420\mu\mu$. For the excited crystal, the
two bands for the c-direction were unchanged, but in the a–b-
plane the absorption had increased greatly, and a new band had
formed in the visible green-yellow, the cause of the phototropic
coloration. Solutions showed no actual bands, but two shoulders
in the ultra-violet in the places of the two crystal bands.

The crystal molecules appear therefore to be identical with
those in solution, but the *parallel orientation* in the crystal allows
the absorption of individual atom-groups to be separately ex-
amined for definite vibration directions of the *electric vector* of
plane polarized light.

Now the maximum excitation (coloration) of the β-tetra-chlor-
α-keto-naphthalin [by white light] is effected only when plane
polarized light falls on the crystal, so that the electric vector
is in the c-direction. This is true, although the absorption is
greater in the a–b-plane. The contradiction to the Grotthus
law is due to the fact that the discoloration by plane polarized
yellow light is effected only when the electric vector is vibrating
in the a–b-plane. The importance of *orientation* for phototropic
[and photochemical] action in solids is very noticeable here.

Weigert points out the fact that the coloration is observed
only in the crystals, not in the solutions, and shows that the
ordered orientation of the molecules in close packing is the con-
dition for the effect. The insolation of the crystal produces
inter-molecular optical influences, which are observed as colora-
tion. This he regards as explainable only by neighboring atom

groups coming nearer together, a possibility in agreement with his more general hypothesis, that the primary photochemical process consists in a separation of those atoms from one another, between which the absorbing electron vibrates. He supposes *this* separation to occur on absorption in the individual molecules which expand and thus brings neighboring molecules nearer together. By absorption in the new intermolecular band these atom groups are again pushed apart, and the initial condition restored.

It is not doubted that Weigert's explanation is correct in main outline. But it may be possible to develop it somewhat in detail, and indicate its extension to other substances than the one he considered. The possibility of this appears to lie, on the one hand, in certain recent developments of valency theory (23), on the other, in the newer theories of ionic deformation, particularly in crystal lattices (24).

Apparently all the phototropic substances contain double bonds or their aromatic equivalents. Double bonds as such are not sufficient for color, i.e., absorption in the visible spectrum, nor in the case of phototropic bodies can an inter-molecular tautomerism (by light) to a quinonoid configuration be invoked, because no effect is produced in solution. The following suggestions do not answer the question, why phototropy is not produced in a much greater number of crystalline organic bodies of the aromatic and heterocyclic series. This question is, however, also unanswered by Weigert's explanation of phototropy. They do appear to proffer a *sufficient* mechanism for the phenomenon; *when* it should also be a *necessary* consequence, will require further constitutional studies and intensive examinations on the lines of Weigert's work to determine.

It has been suggested by Lowry (21) that "while a single bond may be either a covalency [non-polar linkage] or an electrovalency [polar linkage], *a double bond in organic chemistry usually reacts as if it contained one covalency and one electrovalency.*" For many special developments of this theory, reference must be made to Lowry's papers (21). It is to be noted, however, that Sugden and his coworkers (25) on the new molecular volume constant,

the "parachor" have produced definite evidence for the existence, in compounds in the "resting" or inactivated state, of two kinds of double bonds. One of these, common in carbon compounds (but found also in compounds of other elements), causes an *increase* in the parachor of 23.2 units, and another, present in derivatives of the oxy-acids of sulfur and phosphorus, *lowers* the parachor by 1.6 units. The former is the non-polar double bond, the latter (Lowry's) semi-polar double bond. The contraction of the parachor in the latter case is accompanied by a corresponding distortion of the outer shell of electrons. Sugden considers that "Lowry's hypothesis concerning carbon compounds may therefore be taken to mean that the activation of such compounds consists in the transference of an electron whereby a non-polar double bond is converted into one of the semipolar type," a view, which as he remarks, gives a simple connection between the structures on the activated and non-activated states. It is a commonplace of modern photochemistry, that the primary photochemical event consists in the activation of atoms and molecules (26) a view which the writer has emphasized somewhat earlier.[9] At the same time, photochemical activation is only one of several processes by which activation may be induced, all conceptions of an "active" as contrasted with a "passive" form of the molecule dating back to Arrhenius' explanation (27) of the temperature coefficient of chemical reactions. R. G. W. Norrish (28) has recently indicated the importance of traces of polar substances as catalysts for gaseous reactions, and interpreted this in terms of Lowry's hypothesis. "We may thus regard molecular activation as occasioned by a definite change of configuration or distortion of the molecule, brought about by close association with some polar catalyst. Such a change of configuration must take place with the absorption of energy, and thus the activated molecules will be in a more highly energized state than the resting molecules." In comparing this with Lowry's hypothesis, he says "The analogy between the development of an electrovalence on the one hand and the process of

activation on the other is so complete as to suggest that the two phenomena are identical."[10]

It may therefore be suggested that phototropy is due to a process of activation in which non-polar double bonds are converted to semi-polar bonds. In solution, this process, through tending to occur, is neutralized by collisions of the randomly oriented molecules, which prevent the *"set"* being given which is possible in the crystal. The reasoning of Weigert, according
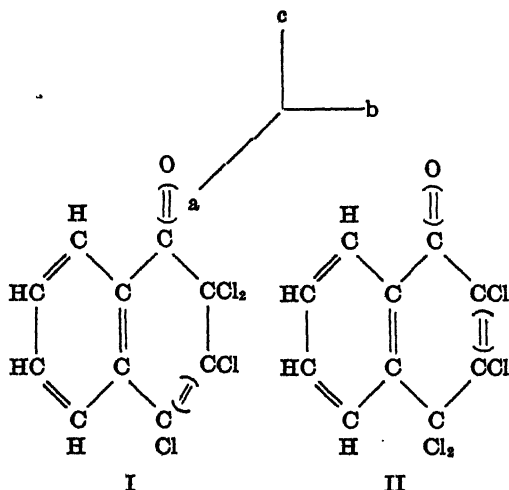
FIG. 4. TWO POSSIBLE CONSTITUTIONAL FORMULAE FOR TETRACHLORKETO-NAPHTHALIN

The vibration planes of valency electrons of the double bonds are indicated by the brackets.

to which the constitution I in figure 4 is to be preferred, from the phototropic phenomena, to constitution II, remains equally valid. The exciting absorption is still to be referred to the C Cl = C Cl linkage, but the essential excitation consists in the change of this to

$$- \overset{+}{C} - \overset{-}{C} -$$
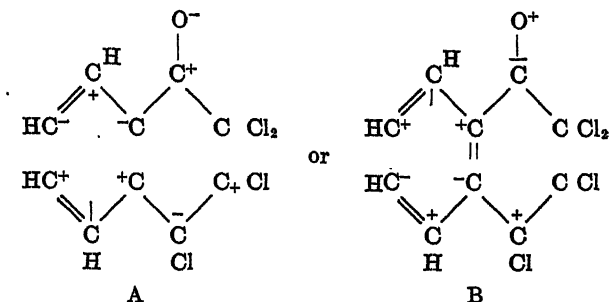$$\quad Cl \quad\; Cl$$
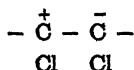
[10] The formation of an electrovalence is not a *phenomenon*, like activation, but the suggested *noumenon*, or actual cause proposed!

As to the stereochemical arrangement of the molecules in the (rhombic) crystal it is suggested by Weigert, since β-tetra chlor α-ketonaphthalene has only one plane of symmetry (in that of the naphthalene nucleus), that the crystal element will have four chemical molecules arranged in some symmetrical fashion in this plane of symmetry. The existence of four chemical molecules in the crystal unit of substituted naphthalene derivatives is confirmed by Sir William Bragg (29), but β-tetra chlor α-ketonaphthalene crystallizes (from benzene) in colorless rhombic bipyramidal crystals, and for this class eight asymmetric molecules are required (30). The molecular symmetry of this highly substituted body is obviously low, lower than that of the mono-substituted naphthalene bodies. In the rhombic bipyramidal class, one molecule is reflected across a plane; the third and fourth are derived from the first and second by reflection across a plane which is at right angles to the first plane; the other four are derived from the first four by reflection across a plane which is at right angles to both the other planes.

We do not know how the atoms in the molecular units of this configuration are arranged in the present case. But it seems reasonable to suppose that they would remain geometrically congruent. The arrangement described would permit in a high degree "tuning" or resonance between the intra-molecular circuits affected by light, i.e., electromagnetically induced, since any two molecules can be brought to coincidence by reflection across a plane of symmetry. Reverting to the structural formula for β-tetra-chlor-α-ketonaphthalene, and accepting Weigert's reasons for constitution I, we can have the transition from the non-polar to the semi-polar double bond, in the C Cl = C Cl when the electric vector is in the c-direction (cf. fig. 4), in both the C Cl = C Cl and the C = C group when it is vibrating in the a and b directions.

It is possible to consider the C = O group as semi-polar in any case, corresponding to its strong chromophoric value. In view of the structure of the crystal, it seems very probable that the alternative semi-polar configurations

$$
\begin{array}{ccc}
& O^- & \\
C^H & \diagup & C^+ \\
HC^- & {}^-C & C\ Cl_2 \\
HC^+ & {}^+C & C_+\ Cl \\
C & & C \\
H & & Cl \\
& A &
\end{array}
\qquad or \qquad
\begin{array}{ccc}
& O^+ & \\
C^H & & C \\
HC^+ & {}^+C & C\ Cl_2 \\
HC^- & {}^-C & C\ Cl \\
C & & C \\
H & & Cl \\
& B &
\end{array}
$$

would be formed as electro-optical images in *pairs*, as reached by reflection across a plane. The electrostatic attractions could produce the further approach of the groups, as supposed by Weigert. In this case, there would be a *contraction* of the chemical molecular volume, since Sugden finds the parachor reduced by 1.6 for a semi-polar bond, *increased* by 23.2 for a double bond. Hence, in this case, a contraction of 1.6 should occur for each double bond changed by light absorption.

Since reversal would occur by the mutual neutralization between two paired molecules, excitation in individual molecules, a possible explanation of Padoa's finding for the bimolecular character of the reverse reaction is afforded.[11]

Attention should be drawn in connection with this to B. H. Wilsden's theory of chemical affinity and electronic structure (31). According to this, bonding by a *magnetic* field gives a non-polar type, by an *electric* field, a polar type, of valency. A change from a non-polar to a semi-polar double bond, according to this, would involve conversion of magnetic to electric energy. Since, in a plane polarized ray, the electric and magnetic vectors are at right angles to each other, Weigert's reasoning on the relation of the C = C and C Cl = C Cl groups, which is based solely on the electric vector, may require reconsideration.

While other phototropic bodies will not necessarily give rhombic

[11] This can also be expressed by the splitting of the half double bond being mono-molecular, the reformation being between + and − charges as bi-molecular. The suggestion made is in agreement with the theory of "induced alternating polarities" of A. Lapworth, *J. Chem. Soc.* 121, 416 (1922); W. Kermack and R. Robinson, *ibid.*, 427.

bipyramidal crystals, they all appear to contain double bonds, and in any likely crystal system, as rhombic pyramidal or mono-clinic prismatic, conditions for electric imaging and coupling will exist, which are not present in solutions. The question obviously arises, why are not all organic substances which contain double bonds and which crystallize, phototropic? Probably a greater number are so than have yet been observed. But this particular activation of double bonds may require certain con-stitutional "weights" or factors not yet ascertained. A further difficulty for the hypothesis that a non-polar double bond is converted to a semi-polar one, is that the reverse reaction might be supposed to necessitate luminescence, which is not observed. This may be due to the fact that relatively few activated bonds are involved, or that the return process is analogous to the radiationless energy transfers of the second kind (32) with excited molecules.

It is evident that on the hypothesis discussed, phototropy in solid phases should have relations with certain valency changes in solutions. The phenomena of *halochromism*[12] are examples, and it is interesting to note that Dietzel and Naton (33) find that *bisdiphenylene fulgide, diphenyl piperonyl fulgide* and di-pipiperonyl fulgide give pale yellow solutions in chloroform or acetic acid, whereas the solutions in mono- di- and tri-chloracetic acids are much darker. Absorption measurements indicate that the production of *halochromism* is optically similar, though quan-titatively less marked, than that produced by conversion of a saturated into an unsaturated compound. Madelung (34) has drawn attention to the halochromic effect of light upon the colorless solution of the carbinol base of para fuchsine, or crystal-violet in acid free acetone, pyridine, etc. We have here the transformation of the pseudo (carbinol)-base into a true base, with a displacement of an electron from the carbon atom next the hydroxyl group to the hydroxyl group, giving a positively charged complex and negatively charged hydroxyl

$$- C \cdot OH \rightleftharpoons - \overset{+}{C} + O\overset{=}{H}$$
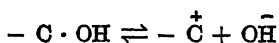
[12] Color formation on salt formation.

The system is unstable and reverts to the pseudo-base, analogously to the phototropic reversal. Again Lifschitz (35) has shown that the *leuco*-cyanides of triphenyl methane dyes are colored in solution by ultra-violet light, both the color and the increased electrolytic conductivity going off again in the dark. This reversion from the true salt to the pseudo-salt is precisely similar to the mechanism suggested for phototropy. The similar thermochromic phenomena, which are also shown by the "free radical" doublets, should be investigated from the viewpoint of photochemical antagonism, the so-called equilibria observed being stationary states depending upon the proportions of radiation incident.

PHOSPHORESCENCE

The striking phenomena of radiation antagonism in relation to phosphorescence have been discussed fully by Lenard (36) and others. Here it may be remarked only that phosphorescence is also a phenomenon of the solid crystalline state (37), and that while a pure substance, such as calcium tungstate, may give only fluorescence in x-rays, the presence of traces of impurities degrades much of the luminescence to phosphorescence. The probability that in this case an electron is first caused to pass over into a new orbit, and that the luminescence is due to its return to its former equilibrium condition does not fully explain the "quenching" action of longer wave-length radiation (38). From a study of the growth and relation of fluorescence and phosphorescence under x-rays, of calcium tungstate, the writer is inclined to regard the completeness of orientation of the atoms in the crystal lattice, and the degree of deformation as of great importance. It will be shown elsewhere that not only is the fluorescent intensity increased as the size and symmetry of the crystals are increased, i.e., the ratio of completely oriented to imperfectly oriented atoms increased, but the introduction of definite groups into the lattice can:

α greatly increase phosphorescence, i.e., lag of re-emission

β neutralize "lag" produced by phosphorogenic groups

The phenomena resemble strikingly many of those observed in the photographic behavior of the silver halides.

That phosphorescence and the photo-electric effect are closely connected is the conclusion of Lenard.[6] In this case, the actual freeing and emission of electrons is understood. One type of photo-electric effect, viz. photo-electric conductivity, has recently been studied under exceptional conditions, in single crystals, by Gudden and Pohl (39). From their experiments they conclude that distinction must be drawn between primary and secondary photo-electric currents. The primary current, they consider, is due to photolytically freed electrons, the secondary, to diminution of specific resistance, e.g., by lattice disorientation (40) or by "coherer effects" at boundary surfaces. The characteristics of the primary current are a finite initial value and saturation proportional to the light energy. Saturation is reached only in single perfect crystals, and *the rate depends upon the intensity*.

In regard to the spectral distribution of the photo-electric current, they say (41) "In consequence, of our experiments we have been more and more led to the conclusion, that inner photo-electric effects in general broaden the optical absorption region toward the longer wave-lengths." This is analogous to the Becquerel effect in photography to be noted shortly.

Crystals in this condition they term "excited," and distinguish between two groups of crystals, in regard to the photocon-ductance. The first group consists of crystals such as *diamond*, and *selenium*, where the absorption is proper to the substance, the second, like NaCl colored by radium or x-rays, have an absorption due to foreign materials. In the case of "foreign" absorption, the total absorption is largely unaltered by excita-tion; in the case of "proper" absorption, it is increased. By a "center" they understand the seat of a quantized, electron split-ting light absorption; these "centers" being always atoms or molecules in some way distinguished from their environment and present only in low concentration.

Generally, a crystal remains "excited" a limited time only. Heat motion gradually reduces it, and irradiation *by wave-lengths for which absorption is first brought about by the excitation with*

shorter wave-lengths and generally termed "long wave radiation" accelerates the return to the initial state. This phenomenon is again an example of radiation antagonism, which is analogous to *phototropy* and to the *Herschell effect* in photography (vide p. 347).

Gudden and Pohl consider this "excitation" process the broad basis for photo-luminescence, light emission by return of a replacement electron (according to Lenard) being only a special case.

These photo-conductance phenomena undoubtedly throw a new light on the same fundamental "inner photo-electric effects" which are variously developed or revealed in *phototropy, phosphorescence*, and *photochemical change*.

ANTAGONISTIC RADIATIONS IN PHOTOGRAPHY

The field of photographic chemistry is rich in asserted examples of antagonistic reactions from different radiations. The difficulty consists in distinguishing pure effects which are definite and reproducible. In fact, the statements in the literature differ widely, and are often in direct contradiction. What follows is less an attempt to summarize the situation and to reach definite conclusions, than to select the outstanding problems and sketch the conditions necessary for isolation of approximately unit effects. The related phenomena which anastomose here may be tabulated as follows:

1. Mechanism of formation of the normal (developable) latent image
2. Photographic sensitization
3. Optical sensitization
4. Desensitization
5. Optical desensitization
6. The Herschell effect (Villard effect)
7. The Becquerel effect
8. Reversal (solarization) by over-exposure
9. Clayden, Wood, etc., reversal
10. Reciprocity and intermittency failures

It is obviously out of the question to discuss all these fully, yet it is practically impossible to deal with those items which appear as antagonistic reactions without referring to the others.

What is essential to bear in mind here is that we are dealing with a probably definite photochemical action (the decomposition of the silver halides into silver and halogen) which is very minute,—the primary reaction—and which is multiplied and made measurable by an independent chemical reaction, development (42). Hence, apparent antagonistic effects of radiation may be either true photochemical antagonisms, like phototropy, and the photochemical equilibria proper, or may be due to disturbances of the product of the primary reaction which interfere with its function in development.

1. Mechanism of formation of the normal developable or latent image

It is fairly well agreed at the present time that the latent image consists of nuclei, chiefly consisting of colloid silver, and distributed as dispersed "development centers" over the silver halide grain (43). There is further good evidence in the case of high speed negative emulsions that these nuclei are formed at pre-existing "sensitivity centers," consisting of some other material than silver bromide, and which are oxidized away by chromic acid (44). The writer has shown that in all probability these "sensitivity centers" consist of silver sulfide (45) and that "oversize" sensitivity centers of silver sulfide produce spontaneous developability (46). In conjunction with Trivelli and Loveland (47), he has proposed an orientation theory of sensitization and latent image formation according to which the silver sulfide nuclei are surrounded by halos of deformed ions in the silver halide lattice, such that the radiation incident on the grain is oriented toward the centers, and the photochemical reduction of the silver halide takes place in their immediate neighborhood. A center makes a grain developable when it reaches a certain size, and the orienting effect of a center is supposed to grow with the formation of silver, i.e., to be auto-catalytic (48). On this theory, the greater sensitivity of the larger grains in the same emulsion is a consequence both of the increased chance of a larger grain having a larger sulfide nucleus, and of the increased mass of silver halide available to afford oriented photo-product.

2. Photographic sensitization

It is evident from the foregoing that sensitization (for light of the same wave-length) and the mechanism of latent image formation are closely connected, so that a complete explanation of one involves the other. While the specific photographic sensitizing by formation of silver sulfide nuclei may be regarded as a demonstrable fact, the orientation theory of sensitization and latent image formation is only a working hypothesis. Alternative views, which accept the silver sulfide nuclei as sensitivity centers, are the following. First, *chemical sensitization.* The silver sulfide nuclei may act as halogen acceptors for the photochemically decomposed silver halide in their immediate vicinity. This would have only a small sensitizing effect, and does not explain why pronounced halogen acceptors in the emulsion do not increase photographic sensitivity.[13]

Second, there may be no *oriented* growth of the sensitivity nuclei on exposure but only a chance distribution of reduced silver atoms to form a nucleus large enough to induce developability. Here again, the measure of sensitization which can be effected, which in favorable cases amounts to 100:1 seems to be out of order with this hypothesis. Direct experimental tests between these hypotheses may not be possible, so that indirect evidence must play a large part in deciding between them, or at least in evaluating their relative importance. It is at least evident that in endeavoring to explain photographic phenomena, what may be termed *topochemical* factors (49) affecting the after-process of development must be given equal consideration with pure photochemical factors. Wherein consists the fundamental photochemical change? It is generally accepted that the latent image consists of colloid (metallic) silver. Further, the process of formation of the silver is considered to consist in the photochemically activated transfer of an electron from a bromide ion to a silver ion (50), according to the equation
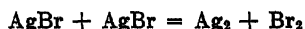
$$Br - \theta \rightarrow Br$$

$$Ag^+ + \theta \rightarrow Ag$$

---
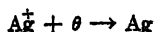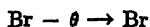
[13] Recently K. C. D. Hickman (*Phot. J.* **51**, 34 (1927) has developed a more promising halogen acceptor theory of this action.

It is to be remarked, however, that such a process leads conceivably to the intermediate formation of the *homopolar* molecule Ag Br, which only decomposes to silver and bromine by the further reaction

$$AgBr + AgBr = Ag_2 + Br_2$$

We may write the reaction then reversibly as

$$2Ag^+ + 2Br^- \rightleftarrows 2AgBr \rightleftarrows Ag_2 + Br_2$$

or more generally

$$2Ag^+ + 2X^- \rightleftarrows 2AgX \rightleftarrows Ag_2 + X_2$$

Where X = Cl, Br or I

Then for given wave-lengths, the equilibrium will lie more to the left or right according to the nature of the halogen X. There is much evidence in favor of the view that with silver iodide, for example, the equilibrium in light lies further to the left than with silver bromide.

3. *Optical sensitization*

The silver halides are normally photosensitive chiefly in their own absorption region in the blue-violet. Sensitivity to longer waves can, however, be increased by various processes of so-called optical sensitization. The best known of these is the use of certain groups of dyes, which sensitize the silver halide for an extended spectral region which, while not identical with the absorption spectrum of the dye in ordinary solvents, is conditioned by this, and is probably identical with the absorption of dye: silver halide combination. But besides this optical sensitizing by dyes, a number of interesting optical sensitizations have been effected by inorganic substances. The first of these was Lüppo-Cramer's demonstration of panchromatic sensitizing by colloid silver (49). The fact that sensitization to long wavelengths could be brought about by this, explained Becquerel's "continuing action" of light, according to which a pre-exposure to blue light would be strengthened by after-exposure to yellow and red rays, an exposure which, without the pre-exposure, would be

ineffective. The colloid silver formed in the first exposure is acting as a sensitizer for the longer waves.

Allied to this is Capstaff and Bullock's discovery (51) that bathing plates in bisulfite, then in frequent changes of slightly alkaline water, confers sensitivity to the longer waves. This may be due to formation of colloidal silver (52) but is also possibly a consequence of adsorption of hydroxyl ions, as noted later.

Actual formation of colloid silver can hardly be the cause of red sensitizing by bathing silver bromide plates in very dilute potassium iodide and cyanide solutions, as discovered by Renwick (53) and investigated in detail by the writer (54). To this collection of inorganic color sensitizing effects must be added a recent discovery of E. R. Bullock that sensitivity in the extreme red (700 mμ +) is induced by treatment with sodium thiosulfate. That this is caused by silver sulfide formation seems to be confirmed by the fact that R. Loveland, also the writer and E. P. Wightman have obtained similar results using substituted thioureas to produce blue-violet photographic sensitivity according to the discovery of the writer. How are these effects related to optical sensitization by dyes, and what is their common denominator? In papers by Fajans and Frankenburger (55) on the influence of ionic adsorption on the photochemical decomposition of the silver halides, conceptions were advanced which open up a new view of these optical sensitizing effects.

They suggested that adsorption of simple cations, as Ag ions, is limited to an electrostatic monatomic layer. The work required, hν, for transfer of an electron from a bromide ion to an *adsorbed* Ag$^+$ ion is less than in the case of the normal surface of the lattice.

The considerations advanced by Fajans and Frankenburger do not, however, seem entirely adequate, for the following reasons. First, the actual surfaces developed in silver bromide crystals are not of the chessboard type described, with alternate Ag$^+$ and X$^-$ ions. This would be a (100) cubic surface, whereas the faces which are found are dominantly octahedral, i.e., all Ag$^+$ ions or all Br$^-$ ions (52). Hence the figure and explanation advanced by Fajans and Frankenburger are not realized, except

perhaps to some extent for adsorption of Ag^+ ions. But if we consider less an adsorption than an inbuilding of foreign nuclei in the crystal grating, then the quantum change, i.e., diminution of $h\nu$ necessary for the reaction $Br^- - \ominus \rightarrow Br$, etc., becomes possible in *virtue of the deformation of contiguous ions of the crystal lattice.*[14] We have here an explanation of the optical sensitizing presented by the very dissimilar substances, *viz.* metallic silver, silver sulfide, silver iodide, and silver cyanide. In each case, insoluble foreign nuclei are built into the crystal. The writer has indicated elsewhere how this effect may not only lead to anomalous optical sensitizing effects, but also contribute to the *concentration* of the blue-violet photochemical decomposition about the sensitivity centers, notably of silver sulfide (56).

```
                              +   +   +
                                  ↑
      + - + - + -           + - + - + -
          ↑
      - + - + - +           - + - + - +

      + - + - + -           + - + - + -
           I                     II
     No adsorption         Adsorbed Ag+ ions
```
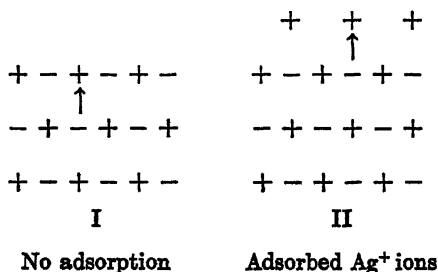
FIG. 5

In normal or dye sensitizing there is probably no such inbuilding, but a surface adsorption of the dye. The principal classes of sensitizing dyes are:

A. Phthaleins, e.g., erythrosin, eosin
B. Cyanins e.g., carbocyanins, isocyanins

The first are acid dyes, forming complex anions. These will be adsorbed chiefly by silver ions, and it is known that these dyes do not sensitize well by bathing, but are assisted by the use of soluble silver salts, i.e., by intermediate silver ions. Going to the pronouncedly basic cyanine dyes (57) these form complex

---

[14] Analogous to the deformations discussed by Fajans (loc. cit.).

cations, and are therefore held by bromide ions.[15]   But, such an adsorption involves reciprocal deformations in the bonded ions, so that the displacement of the spectral sensitizing curve is to be expected.   *Hence, we probably get a superposition of the anomalous and normal optical sensitizing effects.*   The normal effect follows as an inner photo-electric effect in the dye ion, whereby its reduction potential is raised and silver ions are reduced, which form a latent image about "sensitivity specks" just as in the case of the photochemical decomposition of $Ag^+ Br^-$.

4. Desensitization and

5. Optical desensitization

There exist two types of desensitization.   In the first, the desensitizer acts *before* exposure to light, being removed from the plate before exposure.   The most important example is desensitizing by *chromic acid*, the desensitizing action of which was discovered by Lüppo-Cramer (58), and has been fully studied by W. Clark (59) and by Sheppard, Wightman and Trivelli (60).   The action of this, as of permanganate, and probably of iron, copper and uranium salts, when removed from the film before exposure, is to destroy the "sensitivity centers" of silver sulfide by oxidation.

The second type of desensitization is by substances present *during* exposure.   Any of the previously mentioned oxidizers can be used in this way, if not removed before exposure, but the typical desensitizers of the second kind are certain dyes, such as *phenosafranin*, the action of which was also discovered by Lüppo-Cramer (61).   The action of these dyestuffs is greatly reduced, if not entirely eliminated, if they are washed out before exposure (62.)   Hence it appears that they cannot act by oxidizing the the silver sulfide nuclei before exposure but must act during exposure.   Lüppo-Cramer (63) at first suggested that they acted by oxidizing the nascent latent image, even in the presence of

---

[15] The colloidal condition of these dyes in aqueous solutions, while affecting their practical sensitizing powers, is probably of secondary importance for their photochemical activity.

reducing agents, such as alkaline developing solutions. Later he has proposed another theory, according to which they "isolate" the nascent latent image, reducing its capacity to induce development. However, the oxidation theory has been supported by the work of Arens (64) and of Carroll (65). The former regards the action of the dye-desensitizers as an optical sensitizing of the Herschell effect, i.e., of the reversing action of red rays. The latter also supports a similar view. Arens confirmed the
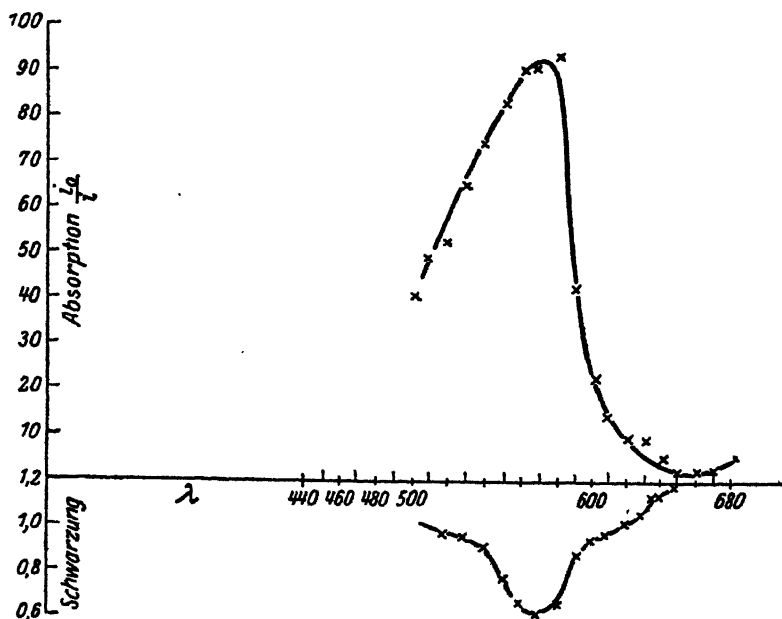

FIG. 6.   ABSORPTION AND DESENSITIZING SPECTRA OF PHENOSAFRANIN

finding of Abney that the presence of free bromide (KBr) increases the Herschell effect, and makes it possible with optically sensitized plates (e.g., with pinacyanol). Halogen absorbing agents, such as $NaNO_2$ do not prevent it so that it seems difficult to attribute it to formation of Br_2. Arens gives the following evidence that the absorption spectrum of phenosafranin coincides with the desensitization spectrum (cf. fig. 6).

Carroll also regards the dye desensitizing action as an oxidation phenomenon, but denies the existence of a specific reversing

effect of red rays (Herschell effect). Concerning the action of the dyes he says. "The dyes, in the dark or in feeble illumination, have an oxidation potential sufficient to prevent formation of a new latent image. On activation by light of their characteristic frequency, their potential is raised sufficiently to cause the destruction of the latent image already existing." The bleaching out of latent image by desensitizing dyes in the presence of free bromide was observed by Lüppo-Cramer and has been confirmed by Dundon in this laboratory.

Before considering this second type of desensitizing as an example of antagonistic action of radiations, it is desirable to notice again optical sensitization. But it will be observed that here also there is a concurrence of topochemical factors (latent image distribution) with pure photochemical equilibria.

While Lüppo-Cramer's suggested topochemical factor, i.e., "isolation" of the latent image nuclei may play a part, the writer is inclined, on the grounds of experiments in progress, to regard the photochemical oxidation potential as the chief factor. Lüppo-Cramer has pointed out that the bleaching out of silver photo-iodide is optically sensitized by dye desensitizers, and that dyes which are *optical sensitizers* for *silver bromide may be optical desensitizers for silver iodide* (61).

Returning to the equations on page 340, we may suppose that dyes are either:

     a. Adsorbed to the Ag^+ ions
     b. Adsorbed to the Br^- ions
     c. Adsorbed to homopolar AgBr (or AgI)

Acid dyes, e.g., erythrosin, giving (complex) anions, would be expected to adsorb to Ag^+ ions, and it is a fact that their sensitizing action is supported by free silver ions.

Basic dyes, e.g., pinachrome, giving complex cations, would be expected to be adsorbed to Br^- ions in the silver halide lattice. It is, however, a noteworthy fact that such dyes are much more soluble (on the alkaline side) in chloroform than in water. *Hence, one may anticipate a possible strong adsorption to homopolar Ag X pairs*, a fact in agreement with the strong adsorption of "basic" dyes to silver iodide. It is also significant that the sensitizing spectra of these dyes with silver bromide approaches more nearly

to the absorption spectrum in chloroform than to that in water or alcohol.[16]

The exact mechanism of optical sensitizing is not yet clear. The relatively simple cases of sensitization by mercury atoms excited to the resonance potential and producing active hydrogen (66) by radiationless collisions have been adduced as significant for photographic sensitizing with the dyes. We have, however, in the dye: silver halide complex, a very complicated system, in which we can scarcely picture activated dye molecules, as wholes, colliding with the silver salt. We must rather conceive of a number of changes being excited by absorption of light, ranging from a transition from non-polar to polar linkages (as in photo-trophy) to reversible electron transfers (reduction and oxidation) and finally to non-reversible changes (hydrogenation and de-hydrogenation). If this series is borne in mind, representable by three energy levels of disturbance of the dye: silver halide system,
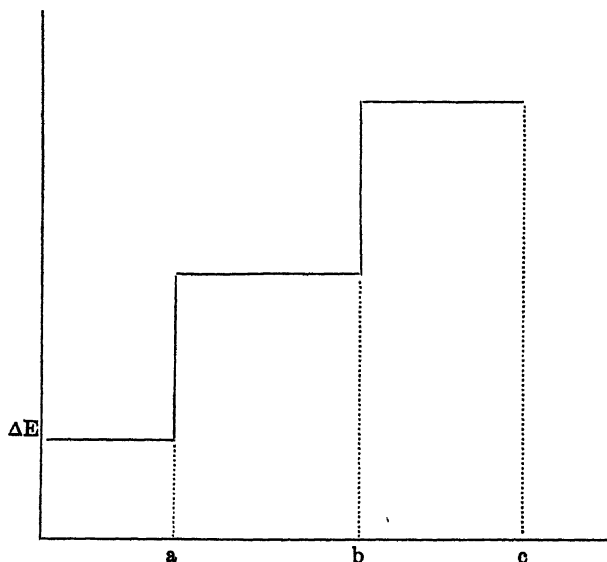
a phototropic changes, *b* inner photoelectric effects, *c* irreversible reductions and oxidations

[16] Cf. S. E. Sheppard, *Phot. J.* **48**, 300 (1908). It is possible that chloroform should be regarded as semi-(bi) polar, rather than completely homopolar.

then the radiation antagonism in dye sensitizing and desensitizing effects appears as a combination of the true (virtual) photochemical equilibria with pseudoantagonistic reactions (destruction of photo-product by certain radiations, as in photocatalyzed autoxidations) and with secondary topochemical effects (Lüppo-Cramer's nucleus isolation).

6. The Herschell effect

This effect, already mentioned, consists in the reversing action of red and infra-red rays in regard to the action of blue and violet. Its existence has been denied by some (65) but appears definite.

Arens (67) regards the Herschell effect as the prototype of optical desensitizing, and as conditioned by the simultaneous presence of Ag Br: Ag: and elementary bromine *or* Br⁻ ions. Since halogen acceptors, e.g., nitrite do not hold it up, but even intensify it, any action of bromine can be regarded only as due to a *transient photochemically active Br from Br⁻ ions occurring in an adsorption layer*. Much more study is required before this effect can be properly interpreted. Its probable connection with the analogous phenomenon in photo-conductivity has already been pointed out (p. 337).

7. The Becquerel effect

The Becquerel effect (68) in photography consists essentially in the sensitizing effect for longer waves of a previous exposure to shorter waves. Lüppo-Cramer has shown that this may be effectively attributed to formation of colloid silver, the latter acting as a panchromatic sensitizer. In line with previous discussion, we may consider this effect, therefore, as a combination of phototropic deformation of the silver halide lattice autocatalytically oriented and accelerated by silver nucleus formation. The phenomenon is, therefore, primarily *photochemical*, and only secondarily *topochemical*. But the complication by topochemical factors is shown by the fact that exposure to x-rays can give a latent image developable by visible light (if halogen acceptors are

present). It stands, therefore, in a reciprocal relation to the Herschell effect.

8. Reversal or solarization by prolonged exposure

Solarization has long been one of the photographic riddles (69). We shall not deal with it in detail, but only note that recent investigations strongly suggested that it is chiefly a *topochemical* effect, i.e., a slowing of development rate due to redistribution of the latent image, i.e., enlargement and anastomosing of the nuclei produced by light (70). Only after fuller investigation in mono-chromatic light of differing wave-lengths will it be possible to discount any photochemical effects, of the pseudo-antagonistic type.

9. Imperfect photochemical integration

The existence of non-integrative effects of photographically effective energy forms is apparent in the *intermittency* and *reciprocity* failures for ordinary light, and more marked in the so-called "Wood series" of "anomalous photographic integrations (71)."

I	II	III	IV
$<S_a + b$	$<S_a + b$	$<S_a + b$	$<S_a + b$
$Sb + a \quad >S_b$	$Sb + a \quad S_b$	$S^b + a \quad >S_b$	$Sb + a \quad \gtrless S_a$
$>S_a$	$>S_a$	$<S_a$	$<S_b$
Incomplete summation	Partial reversals		Total reversal
	(After- (Pre-		
	expo- expo-		
	sure) sure)		

Here S_b denotes the first action alone, S_a the second action alone, $S_b + a$ the two effects together in the order b + a, $S_a + b$ the two effects together in the reverse order a + b; $- S$ always being photographic effect (density).

> b and a may be: Light : X-rays
> Light : spark light
> Light : pressure

10. Intermittency and reciprocity failures

The intermittency and reciprocity failures (72) are special cases of the foregoing incomplete summation. They denote ⌐

a dependence of the formation of the latent developable image upon the rate of supply of energy.

The most important results in this field are those of L. A. Jones and V. C. Hall (73).

Briefly, they find that for the reciprocity failure:

a. The reciprocity failure is approximately symmetrical for low and high intensities. There exists an *optimum* intensity of light energy utilization.

b. The optimum region is wider, or the failure less, the higher the sensitivity of the material.

c. The failure at low intensities involves an actual lowering of D_{max} — (maximum density). That is, the total number of grains made developable for any *time* is a function of the intensity.

While the high intensity failure *may* be a kind of solarization, i.e., a topochemical development effect, the low intensity failure is almost certainly not. It appears that the reverse reaction must here accomplish itself the *better*, the less sensitivity, i.e., the less the presence of sensitizing nuclei.

At present, the writer is inclined to consider this a consequence of lack of orientation of the primary photochemical activation. In absence of sensitivity nuclei, at low intensities, only single, Br^-, Ag^+ ions are affected, and relax from the intermediate stage, whereas in the presence of nuclei ($Ag_2S : Ag$) the process is stabilized at the interface. At high intensities, the radiation density is sufficient for a nucleus large enough to induce sensitization (auto-sensitizing) to be formed (47). The quantum theory of light transmission is not essential, but seems more in consonance with the phenomena.

To sum up, the antagonistic effects of radiations in the photographic process appears to be due:

a. To true photochemical equilibria in the reaction

$$(Ag\ X)_{2m} \rightleftarrows mAg_2 + mX_2$$
$$(AgX)_z$$
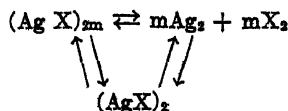
wherein the homopolar bodies play a part as an intermediate "photo-product." Characteristic differences appear in passing

from bromide to iodide, agreeing with the view that the electrons in an I_2 molecule "are hardly more constrained that those in the I^-ion" (74). The equilibrium $(Ag^+ X^-)_{2m} \rightleftarrows (AgX)_{2m}$ is primarily phototropic.

b. To pseudo-antagonistic, or topochemical effects, brought about by the influence of nuclei (of silver sulfide and of silver) upon the subsequent reactions in exposure and development. The photochemical and topochemical effects are "loosely coupled" so that sometimes the progressive, sometimes the regressive actions are assisted thereby.

The conception of an intermediate phototropic equilibrium between heteropolar $Ag^+ X^-$ and homopolar Ag X, which can be further *sensitized* optically by nascent Ag_2 (metal), offers a new interpretation of the schemes advanced by A. P. H. Trivelli (75) and M. Volmer and K. Schaum (71) for the progressive and regressive processes in silver halide emulsions. It remains, however, at present a working hypothesis, to be tested further.

CONCLUSION

The object of this paper is to emphasize the importance of the concept of radiation antagonism for practical photochemistry, and to show that this early induction remains a valuable guide in investigations. In many ways, the concept resembles that of ionic antagonism in bio- and colloid chemistry. This conception also has proven fruitful in leading to new discoveries, and has gradually been stripped of vagueness and associated with the definite mathematical theories of the Donnan membrane equilibrium. Whether a modified or elaborated form of the Perrin-Lewis equation can be brought into similar relation to radiation antagonism remains to be seen. In any case, in the purely physico-chemical phenomena as in the bio-chemical and bio-physical phenomena of antagonism and adaptation, we encounter a class of transformations and equilibria which suggest the necessity of a concept of higher order than those of energy and entropy. The inventor (Johnson Stoney) of the word "electron" for the unit electric charge has proposed for this concept the term "synergy." It may be suggested that in this order of things

(synergic phenomena) we cannot regard a transformation as determined solely by initial conditions, but as co-operatively determined by *certain* initial conditions and final states, which stand to each other in a relation of reciprocity or correspondence. Synergy is apparently an inherent condition of the quantum dynamics of the atom, particularly in respect of light emission and absorption.   Sommerfeld (76) has the following to say concerning the remarkable numerical or arithmetical order of the intensities of spectral lines inside a multiplette, i.e., the values which correspond to the probabilities of transitions from an initial to a final condition.

What is very remarkable in these intensity rules is the reciprocity of initial and end states.   It appears as though what happens is not given by a probability for the initial condition of the atom and a probability for the transition to the end state, but as though initial and end-state determined what happens co-equally according to their corresponding quantum-weight.   This would rather contradict our adopted intuitions of causality, according to which we should prefer to think the development of the process already fixed by the initial state.   It does not seem impossible (ausgeschlossen) to me, that quantum experience may transform our conceptions in this respect.   It has often been observed that for the Bohr emission condition the atom must know beforehand in what state it will finally go before it can radiate.   In the principle of least action also we adopt a teleological, not a causal standpoint.   Such a teleological inversion of the causality principle seems to me to contradict the quantum theory less than it does the classical theory.

At the beginning it was pointed out that antagonistic and coöperative actions of different radiations were of importance for practical photochemistry.   In conclusion, it is suggested that the concept of synergy is inherently important for theoretical photochemistry.

REFERENCES

(1) RUSSELL: *Proc. Roy. Soc. (London)* **61**, 424–433 (1897); **64**, 409–419 (1899).
(2) SHEPPARD AND WIGHTMAN: *J. Frank. Inst.* **195**, 337–347 (1923).
(3) KORNFELD: *Z. wiss. Phot.* **21**, 66 (1921).
(4) RABEL: *Z. wiss. Phot.* **19**, 69 (1919).
(5) RITTER: Beiträge zür nähere Kenntniss der Galvanismus, Vol. 2, p. 282 (1802).

(6) GROTTHUS: *Gilbert's Ann.* **61**, 50 (1819); also W. Ostwald's *Klassiker,* No. 152, p. 94.

(7) BECQUEREL: La Lumière, Ses Causes et Ses Effets (Paris, 1872).

(8) EDER: Hdbuch. d. Photog., Vol. III.

(9) PLOTNIKOW: Lehrb. d. Photochem., p. 53 (Berlin, 1920).

(10) *Ibid.*, p. 258.

(11) PERRIN: *Ann. de Physique* **11**, 108 (1919).

(12) TOLMAN: *J. Amer. Chem. Soc.* **47**, 1524 (1925); Lewis and Smith, *ibid.*, **47**, 1508 (1925).

(13) FRANCK: Address at the 88th Meeting of the Ges. Deutsch. Naturforsch. und Aerzter. Innsbruck, 1924, *Naturwissenschaften,* **12**, 1063 (1924).

(14) BOHR, KRAMERS AND SLATOR: *Phil. Mag.* **47**, 785 (1924).

(15) SHEPPARD, S. E.: Photochemistry.

(16) MARCKWALD: *Z. physik. Chem.* **30**, 143 (1899).

(17) STOBBE: *Ann.* **359**, 1 (1908).

(18) GALLAGHER: *Bull. soc. chim.* IV, **29**, 683 (1921).

(19) SENIER AND SHEPHEARD: *J. Chem. Soc.* **95**, 218, 441 (1909).

(20) PADOA AND COWORKERS: *Atti. accad. Lincei.* (5) **18**, 694 (1909), et seq.

(21) LOWRY: *J. Chem. Soc.* **123**, 822 (1923); *Bull. soc. chim.* **39**, 203 (1926).

(22) WEIGERT: *Z. Elektrochem.* **24**, 222 (1918).

(23) LOWRY: *Trans. Faraday Soc.* **18**, 285 (1923).

(24) FAJANS: *Naturwissenschaften* **11**, 165 (1923).

(25) SUGDEN: *J. Chem. Soc.* **125**, 1177 (1924). SUGDEN, REED AND WILKINS: *Ibid.*, **127**, 1525 (1925).

(26) Symposium on Photochemical Reactions in Gases, Faraday Society (1925).

(27) ARRHENIUS: *Z. physik. Chem.* **4**, 226 (1889); **28**, 317 (1899).

(28) NORRISH: *J. Chem. Soc.* **124**, 3006 (1923).

(29) BRAGG: *Proc. Phys. Soc. London* **34**, 33 (1921).

(30) BRAGG: *J. Chem. Soc.* **121**, 2771 (1922).

(31) WILSDEN: *Phil. Mag.* **49**, 354, 900, 1145 (1925).

(32) KLEIN AND ROSSLAND: *Z. Physik.* **4**, 46 (1921).

(33) DIETZEL AND NATON: *Ber.* **58 B**, 1314 (1925).

(34) MADELUNG: *J. prakt. Chem.* **111**, 131 (1925).

(35) LIFSCHITZ: *Ber.* **52**, 1919 (1919). LIFSCHITZ AND JOFFE: *Z. physik. Chem.* **97**, 426 (1921).

(36) Cf. PRINGSHEIM, P.: Fluorescenz und Phosphorescenz im Lichte der Neucren Atomtheorie, Berlin, Springer (1923).

(37) SHEPPARD: *Illum. Eng.* **10**, 178 (1917).

(38) PRINGSHEIM: *Op. cit.*, p. 109.

(39) GUDDEN AND POHL: *Z. Physik.* **6**, 248 (1921); **3**, 98 (1920); **4**, 206 (1921); **17**, 331 (1923); **18**, 199 (1923); **21**, 1 (1924); **30**, 14 (1924); **31**, 651 (1924); **35**, 243 (1925).

(40) Cf. VON HEVESY: *Z. Physik.* **10**, 80, 84 (1922).

(41) GUDDEN AND POHL: *Z. Physik.* **35**, 249 (1925).

(42) THE SVEDBERG: *Phot. J.* **46**, 315 (1922).

(43) BANCROFT, W. D.: The Theory of Photography, General Introduction Faraday Society Symposium. *Trans. Faraday Soc.* **19**, 243 (1923).

(44) Sheppard, Trivelli and Wightman: *Trans. Faraday Soc.* 19, 296 (1923); Clark: *Ibid.*, 309.
(45) Sheppard: *Phot. J.* 49, 380 (1925).
(46) Wightman, Trivelli and Sheppard: *Phot. J.* 49, 134 (1925).
(47) Trivelli, Loveland and Sheppard: *J. Franklin Inst.* 200, 51 (1925).
(48) *Ibid.*, p. 77.
(49) Lüppo-Cramer: Kolloid Chemie und Photographie, 2nd Edit.
(50) Sheppard and Trivelli: *Phot. J.* 61, 402 (1921). Fajans: *Chem. Ztg.* 1921, p. 666.
(51) Capstaff and Bullock: *Brit. J. Phot.* 67, 719 (1920).
(52) Trivelli, A. P. H., and Sheppard, S. E.: The Silver Bromide Grain of Photographic Emulsions, N. Y., Van Nostrand, 1921.
(53) Renwick: *Phot. J.* 61, 12 (1921).
(54) Sheppard: *Phot. J.* 62, 88 (1922).
(55) Fajans and Frankenburger: *Z. Elektrochem.* 28, 499 (1922); *Z. physik. Chem.* 105, 255, 373, 329 (1923).
(56) Holmes, H. N., ed.: Colloid Symposium Monograph, vol. 3, N. Y. Chemical Catalog Co. (1925) p. 76.
(57) Fischer and Scheibe: *J. prakt. Chem.* 100, 86 (1919).
(58) Lüppo-Cramer: *Phot. Mitt.*, 1909, p. 328.
(59) Clark: *Phot. J.* 63, 230 (1923).
(60) Sheppard, Wightman and Trivelli: *J. Franklin Inst.* 196, 653 (1923); *Trans. Faraday Soc.* 19, 296 (1923).
(61) Lüppo-Cramer: Kolloidchemie und Photographie, p. 82 (1921); *Phot. Korr.*, 1921, p. 276.
(62) Lumière and Seyewetz: *Bull. soc. chim.* 37, 700 (1925).
(63) Lüppo-Cramer: Phot. Ind. 417 (1921).
(64) Arens: *Z. physik. Chem.* 114, 379 (1925).
(65) Carroll: *J. Phys. Chem.* 20, 693 (1925).
(66) Franck: *Z. Physik.* 9, 259 (1922). Bales: *Science Progress* 20, 396 (1926).
(67) Arens: *Z. physik. Chem.* 114, 22 (1925).
(68) Trivelli: *Z. wiss. Phot.* 6, 278 (1908).
(69) Sheppard, S. E., and Mees, C. E. K.: Theory of Photographic Process, N. Y., Longmans, 1907.
(70) Scheffers: Z. Physik. (1924); Arens: *Z. physik. Chem.* 114, 22 (1925).
(71) Volmer and Schaum: *Z. wiss. Phot.* 14, 1 (1914).
(72) Sheppard, S. E., and Mees, C. E. K.: Theory of the Photographic Process, N. Y., Longmans, 1907, p. 214.
(73) Jones and Hall: *J. Optical Soc. Am.* 12, 321 (1926); 13, 443 (1926); 14, 223 (1927).
(74) Smyth: *Phil. Mag.* 50, 366 (1925).
(75) Trivelli: *Z. wiss. Phot.* 6, 197, 237, 273 (1908).
(76) Sommerfeld: Grundlagen der Quantum-theorie und des Bohrschen Atommodelles. (88 Versamm. d. Ges. deutsch. Naturforsch. u. arzte (1924)), *Naturwissenschaften*, 12, 1048 (1924).

'

RECENT ADVANCES IN CELLULOSE AND STARCH CHEMISTRY

HARRY LeB. GRAY AND CYRIL J. STAUD

The Kodak Research Laboratories,[1] Eastman Kodak Company, Rochester, New York

INTRODUCTION

Because of the intensive work being done on cellulose at the present time, the multiplicity of new facts and theories which are being adduced, renders comprehensive treatises on the chemistry of this carbohydrate obsolescent, almost as soon as the mechanics of assembly and publication are possible.

CELLULOSE STRUCTURE, HISTORICAL SUMMARY

For the benefit of those whose main interest has been in branches of Organic Chemistry other than that of cellulose, a brief account of the salient facts in connection with the study of the structure of this substance may be of interest before presenting the latest developments. So much has been published that in a review of this type only what appear, in the opinions of the authors, as the most fundamental work may be recounted and indulgence is asked if important researches seem to have been omitted.

As long ago as 1883 Flechsig (1) claimed that cellulose could be entirely converted into glucose. Ost and Wilkening (2) hydrolyzed cotton with sulfuric acid and obtained an almost theoretica yield of glucose, as estimated polarimetrically and by means of Fehling solution. Subsequently Willstätter and Zechmeister (3) showed that cellulose was soluble in 41 per cent hydrochloric acid and that such a solution upon standing resulted in the hydrolysis of cellulose to glucose. The polarimeter and Fehling solution both indicated a 95 per cent yield of glucose.

[1] Communication No. 332.

Since glucose was not quantitatively isolated, the proof that cellulose could be quantitatively hydrolyzed to glucose was open to criticism.

Cunningham (4) found that two such widely different types as cotton and esparto cellulose, the latter known to contain xylan, gave approximately the same optical rotation in solutions prepared according to the method of Willstätter and Zechmeister.

It was later found by E. C. Sherrard and A. W. Froehlke (5) that cotton cellulose, white spruce, Douglas Fir and yellow birch when dissolved in 40 per cent hydrochloric acid gave practically identical yields of glucose as estimated by means of optical rotation.

That cellulose is capable of hydrolysis quantitatively to glucose was firmly established through the work of Monier-Williams (6) who hydrolyzed cellulose with sulfuric acid and actually isolated pure crystalline glucose to the extent of 91 per cent, and the work of Irvine and Soutar (7) done at approximately the same time. The latter workers subjected cellulose to acetolysis with sulfuric acid in the presence of acetic anhydride and acetic acid and hydrolyzed the resulting glucose acetate with a methyl alcohol solution of hydrochloric acid, thereby obtaining α-methyl glucoside. This was, of course, readily converted to glucose. The yield obtained in this way was 85 per cent. Subsequently in 1922, Irvine and Hirst (8) through modification of the acetolysis method and the use of slightly stronger alcoholic HCl, brought the yield of glucose to 95.1 per cent of that demanded by the expression $(C_6H_{10}O_5)_x \rightarrow C_6H_{12}O_6$.

The methylation of cellulose also contributed to the cellulose $\rightarrow$ glucose question. Through the use of this reaction, Denham (9) introduced 25 per cent of methoxyl into cellulose. The product so obtained, when dissolved in 41 per cent hydrochloric acid and hydrolyzed, gave a crystalline sugar which was identified as 2,3,6 trimethyl glucose. And in 1923 Irvine and Hirst (10) demonstrated that 2,3,6 trimethyl glucose was the only sugar formed when cellulose is hydrolyzed to glucose.

The hydrolysis of cellulose quantitatively to glucose and the proof that 2,3,6 trimethyl glucose is the only glucose derivative

obtained by hydrolysis went far to clear the atmosphere in regard to cellulose structure and rid the subsequent speculations of many ambiguities which had existed prior to that time. It enabled the casting out of much theoretical discussion which had previously appeared, such as the formula for cellulose presented by Cross and Bevan (11) about 1900 in which appeared four free hydroxyl groups. The results of esterification studies had for a long time militated against such a view but the evidence presented in summary above made such a formula quite untenable.

The present views of the structure of cellulose, as held by many workers, regard cellulose as represented by the formulation $[(C_6H_{10}O_5)_x]_y$, in which x equals the number of $C_6H_{10}O_5$ groups in the fundamental molecule and y equals the number of $(C_6H_{10}O_5)_x$ groups joined together by association, polymerization or other forces to produce the colloidal particle or micelle. A notable exception is found in Kurt Hess and his co-workers (12) who maintain that the fundamental cellulose molecule is $(C_6H_{10}O_5)_1$. Their contention is based upon polarimetric data from the behaviour of cellulose in cuprammonium solution, and upon cryoscopic studies of the molecular weight of various derivatives. Recent publications by Pringsheim (13) appear to indicate that he is also of this belief.

On the other hand the isolation of a disaccharide from cellulose, namely cellobiose, leads many to the view that the linkage existing in cellobiose is present in cellulose. Cellobiose is obtained by acetolysis of cellulose, with acetic anhydride, acetic acid and sulfuric acid, with subsequent hydrolysis of the cellobiose octaacetate obtained.

Cellobiose octa-acetate was first isolated in the acetolysis of cellulose by Franchimont (14). It was later studied by Maquenne and Goodwin (15) and other workers. It appears to be one of the degradation products of cellulose. The method of formation lends credence to the likelihood that it is a part of the cellulose molecule. The maximum yield of cellobiose octa-acetate thus far obtained is 35 to 40 per cent. This was accomplished by Haworth and Hirst (16). The yield of cellobiose acetate from the acetolysis of cellulose triacetate and from cellulose appears

to be of the same order of magnitude (17). The Kurt Hess school, however, claim that in the acetolysis of cellulose the carbohydrate is first hydrolyzed to glucose and that cellobiose octa-acetate results from the recombination of two glucose molecules followed by acetylation of the resultant disaccharide.

Work by Irvine and Hirst (18) on the methylation of cellulose and subsequent hydrolysis of the ether thus produced led them to announce a formula for cellulose in which the x in the fundamental molecule was given as 3. The cellulose molecule is represented as an anhydro-trisaccharide as given below.

Irvine and Hirst Formula

CH₂OH

CH—O—CH · CH · CH(OH) · CH(OH) · CH
|
CHOH
|
CHOH
|
CH
|
CH—O—CH · CH(OH) · CH(OH) · CH · CH · CH₂OH
|
CH₂OH
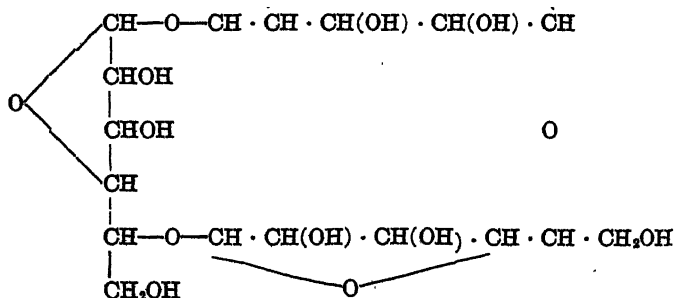
These authors have recently published the results of further work in support of this formula as given in a later paragraph.

Another tool which has proved of value in studying the structure of cellulose has been the use of x-rays. Cellulose and its derivatives yield x-ray spectrograms. Debye and Scherer (19) first recorded the observation that cellulose gave interferences with x-rays very similar to those of crystals, but apparently did not theorize upon their observation. The pioneer work in this field has been done by R. O. Herzog and H. Jancke, who published in 1920 (20) a paper dealing with the constitution of cellulose in which it was stated that $(C_6H_{10}O_5)_2$ represented the fundamental molecule of cellulose. In 1921 Herzog (21) modified

his views and stated that the cellulose unit contained 4 anhydro-glucose units and in 1925 (22) he suggested that the fundamental cellulose molecule might contain 1, 2, or 4 $C_6H_{10}O_5$ groups, but that in contra-distinction to Irvine (18) the value three was excluded by calculations from x-ray data.

With this rather meagre outline, in which an attempt has been made to recount the outstanding features of the work on the structure of cellulose, we pass to the more recent advances in this field, as well as that of the reactions of cellulose and its colloidal nature and behaviour as given in the succeeding paragraphs.

CELLULOSE STRUCTURE, PHYSICO-CHEMICAL RESEARCH

A very important contribution on the structure of cellulose as indicated by x-ray spectra is given in a paper by O. L. Sponsler and W. H. Dore (23). The authors studied the Röntgen diagram of fibers of Ramie and proposed a structure which appears to be consistent with the physical properties and chemical reactions of fibrous cellulose. They consider this substance to be made up of glucose units in the form of amylene oxide rings apparently united by primary valences in chains of indefinite length. These chains are parallel to the longitudinal axis of the fiber and are spaced rectangularly 6.10 x 5.40 Å.u. The ramie fiber is a hollow cylinder in which the crystal units are so placed that one of the diagonals of the 6.10 x 5.40 spacing always occupies a tangential position. The linkage of the chains is alternately from the 1 to 1 and the 4 to 4 carbon atoms of the glucose units. This precludes the presence of the cellobiose linkage in cellulose. They state that a group of eight glucose units is the simplest unit that can represent the structure of cellulose. This corresponds to the crystallographic unit with axes 10.80 x 12.20 x 10.25 Å.u.

The authors consider that the continuous primary valences account for the tensile strength of the fibers in a longitudinal direction while they are stabilized laterally by the secondary valences between the oxygen atoms of adjacent chains. Ester formation is shown to be possible. This may decrease the secondary valence force with a consequent separation of the longitudi-

nal chains and a resultant weakening of the fibrous structure. Sponsler and Dore (23) report a fair agreement with the data obtained by R. O. Herzog (24) but differ in their interpretation. This author (25) while admitting the possibility of the glucose units being glucosidically attached, considers it more likely that a number of $C_6H_{10}O_5$ groups are linked together according to Werner's theory of secondary valences.

E. Ott (26) in discussing the structure of the polysaccharides reports x-ray diagrams in support of the view that the crystallite of cellulose consists of three $C_6H_{10}O_5$ groups and that of lichenin seven. In another paper he (27) declares that cellulose hydrate, oxycellulose and hydrocellulose give diagrams identical with that of lichenin and believes that the crystallite constituents are the same, cellulose being a modification of lichenin. Herzog (28) finds lichenin yields a diagram similar to that of hydrated cellulose, but not identical with it.

Both cellulose nitrate and acetate when produced without destruction of the fibrous form show a crystalline structure according to Herzog (29) and to Ott (30).

Herzog (31) has observed differences in the Röntgen diagram of natural and mercerized fibers and finds that the same differences exist between cotton which has been esterified and deesterified without losing its fibrous form and that which has been in solution. From these data, since it is difficult to conceive that solution would cause a chemical alteration, he considers that mercerization produces a purely physical change. On the other hand, J. R. Katz (32) in a continuation of his previous work (33) adduces further data in support of his contention of the presence of an "alkali cellulose compound." In this instance he worked with alkali dissolved in dilute alcohol. By so doing the level portion of the Vieweg swelling curve disappears, but the same x-ray spectrographic change occurs as is observed in aqueous solutions of approximately the same concentration. When the cotton is washed with dilute alcohol, treated with an acetic acid solution to remove the last traces of sodium and dried, the spectrum of mercerized cellulose is obtained. This view is strengthened by the statement in another paper by Katz (34) that ramie

swollen by zinc chloride or calcium thiocyanate gives an unchanged spectrogram. That cellulose after swelling with HNO_3 and washing shows a Röntgen spectrum characteristic of mercerized cellulose has been shown by J. R. Katz and K. Hess (35). The concentration limits producing this effect are specific gravities 1.38 to 1.42. The so-called Knecht compound, $C_6H_{10}O_5 \cdot HNO_3$ is shown to have a constant composition and an x-ray diagram which differs from that of cellulose.

Work has been done by Herzog and Laski (36) on the absorption spectra of thin films of cellulose and nitrocellulose in the infra red as a possible clue to the structure of these compounds.

Alb. Frey (37) determined the refractive indices of various kinds of cellulose fibers and from the data obtained concluded that cotton cellulose and other types of fibrous cellulose are identical.

CELLULOSE STRUCTURE, CHEMICAL RESEARCH

Hess (12) in continuance of his work in support of the contention that $(C_6H_{10}O_5)$ is the molecular unit of cellulose has recalculated (38) the data given by Herzog (29) and found the number of $C_6H_{10}O_5$ groups per crystal unit to be 1.62 and 3.21 instead of 15.9 and 32.1 as originally recorded by Herzog. From this Hess concludes that the number 4 has nothing to do with the cellulose molecule. Herzog (39) admits the calculation error, but still maintains 2 or 4 $C_6H_{10}O_5$ mols per unit in the cellulose derivatives.

Hess (40) found from freezing point determinations in absence of air that di- and triacetyl cellulose dissolve in glacial acetic acid in concentrations of 0.05 to 0.6 per cent in monomolecular form corresponding to $C_6H_{10}O_5$. He favors the Naegeli micellar hypothesis and considers that glacial acetic acid is able to disperse the molecular building stones of the micelles. The above author (41) gives a review of his results obtained on the study of cellulose.

H. Pringsheim (13) heated cellulose triacetate in naphthalene and in tetrahydronaphthalene at 235°C. for two hours and observed that it is broken down without loss of material and without splitting off acetyl giving glucose anhydride triacetate. Saponi-

fication of this gives a substance which he calls cellosan, acetolysis of which yields cellobiose octa-acetate. Hess (42) in a repetition of this work, finds that acetyl is split off during the heating.

The same author (43) observed in the methylation of glucoses a shifting of the oxygen bridge under certain conditions. In another paper (44) the preparation of crystalline trimethyl cellulose is described. From cryoscopic work in vacuo he found properties similar to those of cellulose acetate under the same conditions. He concludes that methylation is the direct introduc-
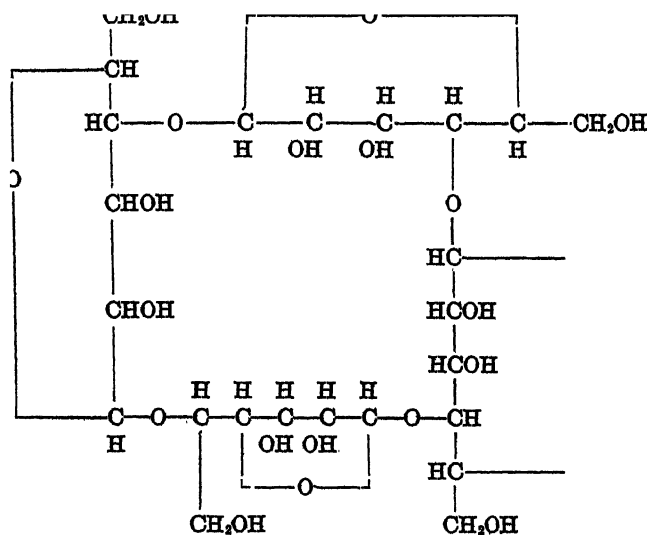

Fig. 1

tion of the methyl group except the possible wandering of the oxygen bridge.

Cellulose degraded by acetolysis and then deacetylated was shown by J. C. Irvine and G. J. Robertson (45) upon methylation and analysis of the products so obtained, to give 35 per cent anhydro-triglucose. They believe that the triglucose unit must be at least one-third of the cellulose aggregate.

A constitutional formula-for cellulose has been published by H. LeB. Gray (46), involving four $C_6H_{10}O_5$ groups or a multiple thereof as indicated in the structure given in figure 1. This

structure is based on the spectrographic data and the chemical evidence that one hydroxyl per C_{24} unit possesses properties differing from the other eleven hydroxyls present. A hypothetical breakdown of such a molecule makes possible the existence of a disaccharide not known at the time of publication as well as cellobiose, celloisobiose and glucose. Upon methylation and hydrolysis it would yield 2,3,6 trimethyl glucose in accordance with Irvine (18).

By the acetolysis of cellulose, Hess and Hermann Friese (47) using the method of Ost with the sulfuric acid decreased to one-tenth, obtained a new derivative which they have determined to be a hexa acetyl biosan. Upon saponification an anhydro disaccharide is produced. Further degradation yields cellobiose and isocellobiose.

Fritz Micheel and Watroslow Reich (48) claim to have isolated from cellulose, by acetylation, deacetylation of the product obtained and reacetylation in pyridine, a new body which they believe may be related to the "Kittsubstanz" proposed by Herzog (49).

DISPERSOIDOLOGY OF CELLULOSE

During the past year the dispersoidology of cellulose and its derivatives has been the subject of many investigations. According to Herzog (50) electrolyte free polysaccharides do not swell in water. He considers that when swelling does take place the crystals break apart and the "Kittsubstanz" is partially destroyed. There are, in his opinion (51), two types of swelling agents; the first includes sodium hydroxide, cuprammonium and nitric acid; the second, salts and their water of crystallization. Under tension with the first type, the oblong crystallites imbedded in the "Kittsubstanz" become liquid. Release of the tension does not induce crystallization but results in coagulation.

W. Gordon (52) assumes the cotton fiber to have the properties of liquid crystals distributed in an amorphous cementing substance.

From the studies of solutions of cellulose in Schweitzer's reagent, Herzog and Krüger (53) concluded that the original

cellulose crystals whose size depended on the nature and previous history of the cellulose material, always dispersed into primary particles of the same size.

This appears to be contradicted by M. Numa (54) who claims that the color intensity of cuprammonium solutions of cellulose is directly related to the degree of dispersity.

Because of controversy concerning previous papers about the ripening of viscose, Herzog (55) published an account of the work which led to his papers on the viscose ripening process. The methods used for determining particle size are given. He states, in one of his papers (56) that the ripening of viscose is a coagulation phenomenon in which secondary particles consisting of series of rod shaped micelles are formed.

An investigation of the viscosity of viscose solutions led E. Berl and A. Lange (57) to conclude that ripening of viscose is not a polymerization effect of the cellulose xanthate molecule but that the cellulose formed by the splitting off of xanthic acid coagulates.

W. Von Neuenstein (58) states that cellulose nitrate and cellulose acetate solutions which have become thin on standing increase in viscosity when stirred. This effect is probably due, in his opinion, to the breaking up of secondary micelles.

Many data concerning cellulose nitrate solutions have been published by J. W. McBain, C. E. Harvey and L. E. Smith (59). They consider the solution of cellulose nitrate to be a direct combination between the solvent and suitable complementary chemical groups in the solute. Loose ramifying aggregates of colloidal particles held by local and specific bonds of residual affinity are believed to account almost entirely for the apparent viscosity of such solutions.

A study of the jellies formed by cellulose triacetate with benzyl alcohol and water has been made by H. J. Poole (60).

P. Karrer (61) observed the rate of zymolysis of cellulose by snail cellulase. He concuded the resistance to zymolysis is decreased by processes which cause a loosening of the micellar structure.

ALKALI CELLULOSE

Work on the action of alkali hydroxides on cellulose has brought forth numerous papers during the past year. The conclusions, in the main, represent those previously arrived at, with slight discrepancies.

Herzog (50) considers that the mass of water imbibed cannot be related to the heat of mercerization. However, Katz (62) believes that the taking up and liberation of water in the treatment of cellulose with alkali of different strengths plays an important part.

Making certain assumptions, W. Gordon (52) has derived a formula from which contraction on mercerization can be calculated.

Liquid ammonia at $-33°$ to $-35°$ according to G. Bernard (63) causes swelling of cotton fiber nearly equal to that produced by NaOH. The action of 22 per cent ammonium hydroxide on raw and wax free cellulose is similar to that of sodium hydroxide, but not as intensive. The products of the action of 22 per cent ammonium hydroxide still show intact external structure of cellulose and show no change in furfural, copper number, swelling power and fat content. If cellulose is treated with 22 per cent ammonia at $200°$ for forty-eight hours, a brown powder is obtained containing 20 per cent of nitrogen.

In the preparation of viscose E. Heuser and Schuster (64) found that lithium, potassium, sodium and rubidium hydroxides form true viscoses. With the first three the ratio of alkali to cellulose is one alkali hydroxide molecule per $(C_6H_{10}O_5)_2$. In the case of rubidium three $C_6H_{10}O_5$ groups are required per molecule of $Rb(OH)$.

Emil Heuser (65) discusses the relationship between degree of swelling of cellulose fibers and concentrations and electro conductivity of aqueous and alcoholic alkaline hydroxide solutions at the point of quantitative formation of alkali cellulose as determined by means of x-ray.

ACIDS ON CELLULOSE

Further work has been done on solutions of cellulose in sulfuric and hydrochloric acids by K. Atsuki and F. Minaki (66) (67). In each case two points of inflection in the rotation-time curves were observed. The viscosity decreases rapidly until the first point of inflection is reached after which it remains nearly constant as the cellulose changes from colloid to crystalloid products.

According to Lieser (68) HCl supersaturated at zero degrees converts cellulose quantitatively into alkali soluble cellulose. Schwalbe's copper number shows no apparent increase in the reducing power of the product. The cellulose degradation cannot be explained, but the author considers that it is not a case of hydrolysis.

VISCOSE

Several interesting papers concerning viscose have appeared. Among these the following may be mentioned.

E. Berl and Johann Bitter (69) by the conversion of cellulose which had been alkylated, so as to contain one to two and one-half alkyl groups per C_6 unit, into viscose, arrive at the conclusion that there must be two free hydroxyl groups in the cellulose molecule for complete xanthation.

A paper by Heuser and Schuster (70) states that $(C_6H_{10}O_5)_2$ is the lowest xanthate formed. This breaks down to give a compound corresponding to $(C_6H_{10}O_5)_3$ and then to one of $(C_6H_{10}O_5)_4$. After this the breakdown is indefinite.

In a series of articles J. d'Ans and A. Jager (71) discuss the ripening of viscose. Experimental work shows that the decrease in the number of xanthate groups is regular and that the rate of separation of xanthate groups increases with rising temperature. The properties of the viscose solution and the number of xanthate groups have a clear relationship.

R. Bernhardt (72) suggests that acetic acid may be used for following the course of ripening of viscose.

The literature of the cellulose derivatives has been increased by the publication of many papers.   A few of the more important contributions are given.

A study of the nitration of various celluloses was made by M. G. Morin (73) in which he makes various suggestions for the use of new materials in the production of cellulose nitrate.

F. Blechta (74) confirms the fact that the instability of gun cotton is due primarily to the presence of sulfuric acid esters of cellulose.   By a new method of stabilization with concentrated nitric acid he has shown that nitro hydrocellulose and nitro oxycellulose are not in themselves unstable.   The prejudicial effect on the stability of crude nitrocellulose of N_2O_3 in the nitric acid has been confirmed.   A short boiling with dilute caustic soda solution renders the nitrous acid esters harmless.

A description of the methods of analysis of cellulose nitrate is given by R. Gabillion (75) (76).

The hydrogen ion method for the determination of the stability of nitro-cellulose as proposed by N. L. Hansen is discussed by L. Metz (77).

From the experiments on the esterification of alkali cellulose with acid chlorides G. Kita et al. (78) arrive at the conclusion that only the chemically combined alkali takes part in the esterification. The same authors (79) treated dried cotton paper with stearyl or palmityl chloride in pyridine for eighteen hours, under reflux and obtained products the analysis of which indicated the formation of tri-esters.

H. Gault and P. Ehrmann (80) by the action of the chlorides of lauric, palmitic and stearic acids on hydrocellulose "Girard" prepared the mono- di- and tri-esters of the respective acids. Esters of fatty acids of cellulose containing more than five carbon atoms are described in a series of papers by the same authors (81).

The naphthenic acid esters of cellulose have been prepared by G. Kita and coworkers (82) (83).   A description of the compounds is given.

Hydrophilic cotton, "Girard" hydrocellulose, cellulose from

cuprammonium, from viscose and alkali cellulose were used. An improvement was made by using $SOCl_2$ instead of the usual PCl_5 or PCl_3 in the preparation of the acid chlorides used.

The viscosity of cuprammonium solutions of cotton cellulose has been investigated by F. C. Hahn and H. Bradshaw (84) who find that higher viscosities are .obtained with linters than with long fibered cotton.

O. Faust (85) has found that artificial silk fibers or films of all types when stretched show double refraction, and lose the property again when released.

ANALYTICAL

A. Kiesel and N. Semigasnovski (86) have published a method for determination of cellulose in plant material, by first hydrolyzing with dilute HCl to remove easily hydrolyzable material and then hydrolyzing the cellulose with sulfuric acid and determining the glucose obtained.

A discussion of the determination of alpha cellulose is given by P. Waentig (87) in which he mentions several unsatisfactory aspects of the procedure and offers recommendations for remedying them.

BIOCHEMICAL

In the investigation of pine wood, J. Marcusson (88) observed during rotting the cellulose is converted to oxidized cellulose and pectins.

S. Winogradski (89) has described the isolation in pure culture of a cellulose digesting vibrio. The gelatinous substance obtained is thought to be a colloidal form of oxycellulose.

A very sensitive method for detecting differences between the various types and varieties of cellulose, through hydrolysis with snail cellulase is described by P. Karrer (61). He has also shown that in the hydrolysis of cellulose by this enzyme the further addition of the same does not carry the action beyond that obtained by the original present. It is the opinion of the author that cellulose consists of two constituents which differ in their behaviour toward enzymes.

In another paper P. Karrer and P. Schubert (90) communicate the results of their study of the rate of enzyme action of snail cellulase on cellulose and viscose filaments. They conclude that the seration of the surface, the character of the surface membrane and possible salt occlusion are important contributory factors in this reaction.

J. A. Viljoen, E. B. Fred, and W. H. Peterson (91) reported the isolation in pure culture of a thermophilic organism which destroys cellulose at 65°. After growth on cellulose free media it is unable to ferment cellulose.

<center>MISCELLANEOUS</center>

That the age of the fiber, whether living or dead, is neglected, though chemical and colloid changes may take place, is noted by C. G. Schwalbe (92). He mentions that standard cellulose should be used in all investigations and refers to the work of the American Chemical Society's Committee.

P. Ehrmann (93) gives a review of the chemistry of cellulose and its derivatives with 123 references.

A review of the structure of the cotton fiber by A. J. Turner (94) has been published with a bibliography of sixty-five references.

<center>STARCH</center>

C. L. Alsberg, E. P. Griffing, and J. Field (95) describe a method of preparing a starch solution, by first grinding in a ball mill followed by sifting 2 per cent of the ground material into distilled water, and stirring for approximately one hour. The liquid is centrifuged at 2000 r.p.m. for one-half to one hour. The supernatant liquid is stored in a bottle, the surface being covered with toluene. Clear solutions are obtained in this manner, which will keep for many months and are considered superior for use as an indicator to Lintner's soluble starch or starch paste.

In a series of communications Kikuo Nagai (96) reported the results of ultra-microscopic studies on the fermentative processes of starch and on the starch-iodine reaction. Also starch digestion with pancreatin.

A study of the acid hydrolysis of starch by D. R. Nanji and R. G. L. Beazeley (97) led to the conclusion that the gelatinization of ordinary starch is due to the calcium salt of the amylophosphoric ester rather than to the ester itself. It was found that soluble starch was able to take up calcium from water when the amount present was as low as 1 part of calcium per 100,000 parts of water.

A triacetyl amylose was obtained by Max Bergmann and Ewald Knehe (98) by the acetylation of amylose in pyridine. Saponified with alcoholic potassium, amylose is regenerated which shows all of the properties of the original material, and when reacetylated in the same manner gives the same triacetate. From this the authors conclude that amylose is a strongly aggregated glucose anhydride and not a polysaccharide.

R. Kuhn and W. Ziese (99) found that upon the degradation of monomethyltrihexosan a 6-methyl-glucose is obtained. The results definitely exclude Pringsheim's (100) formulas for the hexosan. The authors conclude that the view generally held that oxygen bridges occur in starch through the C-6 position must be abandoned.

J. C. Irvine and J. MacDonald (101) found that the exhaustive methylation of starch yielded three products of constant composition and properties; (1) dimethyl starch with 32 per cent methoxyl, (2) methylated starch containing 36.3 per cent methoxyl, and (3) trimethyl starch, 43.7 per cent methoxyl. The last two yielded 2,3,6 trimethyl methyl glucoside, when hydrolyzed with methyl alcohol and hydrochloric acid. The 2,3,4 product which is derived from maltose could not be found.

In a paper on the constitution of starch, A. Pictet (102) shows that a linear relationship exists between the molecular optical rotatory power of degradation products of starch and their coefficients of polymerization. He considers that a condensation of three molecules of hexahexosan, $(C_6H_{10}O_5)_6$, is the simplest molecule of soluble starch. In these molecules, $(C_6H_{10}O_5)_{18}$, all of the atoms are held by their ordinary valences. In starch these molecules are probably associated.

In a continuation of his work on the chemistry of starch, H.

Pringsheim and J. Leibowitz (103) discuss the molecular magnitude and association of polyamyloses. The same author (104) gives a summary of his views on the constitution of cellulose and starch.

REFERENCES

(1) FLECHSIG, E.: Zeit. Physiol. Chem. VII, 523 (1883).
(2) OST, H., AND WILKENING, H.: Chem. Zeit. XXXIV, 461 (1910).
(3) WILLSTÄTTER, R., AND ZECHMEISTER, L.: Ber. 46, 2401 (1913).
(4) CUNNINGHAM, M.: J. Chem. Soc. 113, 173 (1918).
(5) SHERRARD, E. C., AND FROEHLKE, A. W.: J. Am. Chem. Soc. 45, 1729 (1923).
(6) MONIER-WILLIAMS, G. W.: J. Chem. Soc. 119, 803 (1921).
(7) IRVINE, J. C., AND SOUTAR, C. W.: J. Chem. Soc. 117, 1489 (1920).
(8) IRVINE, J. C., AND HIRST, E. L.: J. Chem. Soc. 121, 1585 (1922).
(9) DENHAM, W. S., AND H. WOODHOUSE: J. Chem. Soc. 103, 1735 (1913); 105, 2357 (1914).
(10) IRVINE, J. C., AND HIRST, E. L.: J. Chem. Soc. 123, 518 (1923).
(11) CROSS, C. F., AND BEVAN, E. J.: J. Chem. Soc. 79, 366 (1901).
(12) HESS, KURT AND HIS CO-WORKERS: Ann. 435, 1 (1923).
(13) PRINGSHEIM, H.: Ann. 448, 163 (1925).
(14) FRANCHIMONT, A. P. N.: Ber. 12, 1941 (1879).
(15) MAQUENNE, L., AND GOODWIN, W.: Bull. Soc. Chim. 31, 854 (1904).
(16) HAWORTH, W. M., AND HIRST, E. L.: J. Chem. Soc. 119, 193 (1921).
(17) Unpublished results of the authors.
(18) IRVINE, J. C., AND HIRST, E. L.: J. Chem. Soc. 123, 521 (1923).
(19) DEBYE, P., AND SCHERER, P.: Physik. Z. 17, 277 (1916).
(20) HERZOG, R. O., AND JANCKE, H.: Ber. 53, 2162 (1920).
(21) HERZOG, R. O.: Cellulosechemie 2, 102 (1921).
(22) HERZOG, R. O.: Cellulosechemie 6, 39 (1925).
(23) SPONSLER, O. L., AND DORE, W. H.: Colloid Symposium Monograph, (1926).
(24) HERZOG, R. O.: J. Phys. Chem. 30, 455-67 (1926).
(25) HERZOG, R. O.: Pulp and Paper Mag. Can. 24, 694-9 (1926).
(26) OTT, E.: Physik. Z. 27, 174-177 (1926).
(27) OTT, E.: Helv. Chim. Acta 9, 31-32 (1926).
(28) HERZOG, R. O.: Paper presented before the Dominion Convention of Chemists, Montreal, May 31-June 3, 1926.
(29) HERZOG, R. O.: Helv. Chim. Acta 9, 631-633 (1926).
(30) OTT, E.: Helv. Chim. Acta 9, 378-379 (1926).
(31) HERZOG, R. O.: Ber. 60B, 600-603 (1927).
(32) KATZ, J. R.: Z. Electro. Chem. 32, 125-128 (1926).
(33) KATZ, J. R.: Cellulosechemie 6, 35 (1925).
(34) KATZ, J. R.: Z. physik. Chem. 124, 352-358 (1926).
(35) KATZ, J. R., AND HESS, K.: Z. physik. Chem. 122, 126-136 (1926).
(36) HERZOG, R. O., AND LASKI, G.: Z. physik. Chem. 121, 136-142 (1926).
(37) FREY, ALB.: Kolloidchem. Beihefte 23, 40-50 (1926).

(38) HESS, K. Naturwissenschaften **14**, 822–823 (1926).
(39) HERZOG, R. O.: Naturwissenschaften **14**, 952 (1926).
(40) HESS, K. Naturwissenschaften **14**, 435–436 (1926).
(41) HESS, K. Kolloidchem. Beihefte **23**, 93–108 (1926).
(42) HESS, K. Ann. **450**, 65–74 (1926).
(43) HESS, K. Ann. **449**, 146–155 (1926).
(44) HESS, K. Ann. **450**, 29–40 (1926).
(45) IRVINE, J. C., AND ROBERTSON, G. J.: J. Chem. Soc. **1926**, 1488–1501.
(46) GRAY, H. LeB.: Ind. Eng. Chem. **18**, 811 (1926).
(47) HESS, K., AND FRIESE, HERMANN: Ann. **450**, 40–58 (1926).
(48) MICHEEL, FRITZ, AND REICH, WATROSLOW: Ann. **450**, 59–65 (1926).
(49) HERZOG, R. O.: Naturwissenschaften **12**, 954 (1924).
(50) HERZOG, R. O.: Kolloid Z. **39**, 98–107 (1926).
(51) HERZOG, R. O.: Papier-Fabr. **24**, 226–227 (1926).
(52) GORDON, W.: Kolloid Z. **39**, 107 (1926).
(53) HERZOG, R. O., AND KRÜGER, D.: Kolloid Z. **39**, 250–252 (1926).
(54) NUMA, M.: Cellulosechemie **7**, 136 (1926).
(55) HERZOG, R. O.: Kolloid Z. **39**, 252–262 (1926).
(56) HERZOG, R. O.: Papier-Fabr. **24**, 94–97 (1926).
(57) BERL, E., AND LANGE, A.: Cellulosechemie **7**, 145–147 (1926).
(58) VON NEUENSTEIN, W.: Kolloid Z. **39**, 88–90 (1926).
(59) McBAIN, J. W., HARVEY, C. E., AND SMITH, L. E.: J. Phys. Chem. **30**, 312–352 (1926).
(60) POOLE, H. J.: Trans. Faraday Soc. **22**, 83–101 (1926).
(61) KARRER, P.: Melliand's Textilberichte **7**, 23–24 (1926).
(62) KATZ, J. R.: Z. Elektro. Chem. **32**, 269 (1926).
(63) BERNARD, G. Z.: angew. Chem. **38**, 1195–1197 (1926).
(64) HEUSER, E., AND SCHUSTER, M.: Cellulosechemie **7**, 17–55 (1926).
(65) HEUSER, E.: Cellulosechemie **8**, 31–32 (1927).
(66) ATSUKI, K., AND MINAKI, F.: Cellulose Ind. Tokyo **2**, 3–10 (1926).
(67) ATSUKI, K.: Cellulose Ind. Tokyo **1**, 53–61 (1925).
(68) LIESER, TH.: Cellulosechemie **7**, 85–88 (1926).
(69) BERL, E., AND BITTER, JOHANN: Cellulosechemie **7**, 137–145 (1926).
(70) HEUSER, E., AND SCHUSTER, M.: Cellulosechemie **7**, 17–47 (1926).
(71) D'ANS, J., AND JAGER, A.: Kunstseide **8**, 17–19, 43–46, 57–59, 82–84, 110–111 (1926).
(72) BERNHARDT, R.: Kunstseide **8**, 257–260, 314–319 (1926).
(73) MORIN, M. G.: Mem. poud. **22**, 57–68 (1926).
(74) BLECHTA, F.: Vojinsko-Techenike Zpravy XI–XII, Chemicky obzor **1** (1926).
(75) GABILLION, R.: Rev. mat. plast. **81** (1926).
(76) GABILLION, R.: Farbe und Lack. 559–560, 570–571 (1926.)
(77) METZ, L.: Z. ges. Schiessprengstoff **21**, 186–188 (1926).
(78) KITA, G., ET AL.: Cellulosechemie **7**, 125–133 (1926).
(79) KITA, G., ET AL.: Kunststoffe **16**, 41–43 (1926).
(80) GAULT, H., AND EHRMANN, P.: Bull. soc. Chim. **39**, 873–883 (1926).
(81) GAULT, H., AND EHRMANN, P.: Caoutchouc and Gutta-Percha **23**, 13,312–13,313; 13,888–13,889 (1926), **24**, 13,453–13,455 (1927).

(82) KITA, G., AND CO-WORKERS: Kunststoffe 16, 167–168, 199–201 (1926).
(83) KITA, G., AND CO-WORKERS: Cellulose Ind. Tokyo 2, 31–38 (1926).
(84) HAHN, F. C., AND BRADSHAW, H.: Ind. Eng. Chem. 18, 1259 (1926).
(85) FAUST, O.: Ber. 59B, 2919–2920 (1926).
(86) KIESEL, A., AND SEMIGASNOVSKI, N.: Ber. 60B, 333–338 (1927).   ·
(87) WAENTIG, P.: Tech. wiss. Teil, Papier-Fabr. 24, 869 (1926).
(88) MARCUSSON, J.: J. Soc. Dyers and Color. 42, 355 (1926).
(89) WINOGRADSKI, S.: Compt. rend. 183, 691–694 (1926).
(90) KARRER, P., AND SCHUBERT, P.: Helv. Chim. Acta 9, 893–905 (1926).
(91) VILJOEN, J. A., FRED, E. B., AND PETERSON, W. H.: J. Agri. Sci. 16, 1–17
        (1926).
(92) SCHWALBE, C. G.: Kolloid Z. 39, 178 (1926).
(93) EHRMANN, P.: Caoutchouc and Gutta-Percha 23, 13,030–13,032, P. 5;
        13,064–13,065; 13,099–13,102; 13,138–13,139; 13,175–13,176; 13,240–13,241;
        13,275–13,276 (1926).
(94) TURNER, A. J.: Agr. J. India Vol. 21, 274–294 (1926).
(95) ALSBERG, C. L., GRIFFING, E. P., AND FIELD, J.: J. Am. Chem. Soc. 48,
        1299–1300 (1926).
(96) NAGAI, KIKUO: Acta, Schol. Med. Univ. Imp. Kioto 7, IV, 569–575, 577–
        581, 8, I, 57–73, 75–86.
(97) BEAZELEY, R. G. L.: J. Soc. Chem. Ind. 45, 215–219T (1926).
(98) BERGMANN, MAX, AND KNEHE, EWALD: Ann. 452, 141–150 (1927).
(99) KUHN, R., AND ZIESE, W.: Ber. 59B, 2314–2316 (1926).
(100) PRINGSHEIM, H.: Ber. 57B, 1581–1598 (1924).
(101) IRVINE, J. C., AND MACDONALD, J.: J. Chem. Soc., 1926, 1502–1518.
(102) PICTET, A.: Helv. Chim. Acta, 9, 33–37 (1926).
(103) PRINGSHEIM, H., AND LEIBOWITZ, J.: Ber. 59B, 2058–2064 (1926).
(104) PRINGSHEIM, H.: Z. angew. Chem. 39, 118 (1926).

THE MOLECULAR STRUCTURE OF WATER

HARRIS MARSHALL CHADWELL

Department of Chemistry, Tufts College, Massachusetts

Because of the importance of water in the theoretical specula-
tions as well as the practical routine of chemists, a review of the
theories concerning its structure is pertinent. The physical prop-
erties of liquid water are different from those of most liquids.
As the distinction between normal and associated liquids has de-
veloped, water has been accepted as an excellent example of the
second class. A record of the development of the various theories
of the liquid state, as they have been proposed to explain partic-
ular physical properties, would elucidate the case of water. All
these criteria of association qualitatively classify water as asso-
ciated yet the different quantitative methods no not yield con-
cordant results as to the extent of that association (1). Hence,
we shall primarily consider those theories which deal with water
itself, reviewing the hypotheses that have been advanced con-
cerning the equilibria existing in the liquid state.

"ICE MOLECULES"

Whiting (2) was the first to speak of the possibility of "solid
particles" in liquid water. In a thesis on a "Theory of Cohe-
sion" published in 1884, he developed mathematically a theory
of cohesion based on the action of three pressures, an external
pressure, a pressure caused by the kinetic motion of the particles,
and a pressure due to the affinities of one molecule for another.
He assumed that cohesive forces between two particles depend
upon the fourth power of the distance between them. From
these assumptions he derived formulas connecting the physical
properties of a liquid, such as volume change with temperature
and pressure, latent heat, specific heat, critical phenomena, etc.
Because the data for water did not satisfy his equations which

were applicable to most liquids, he came to the conclusion that only those liquids which suffered no molecular rearrangements when heated or put under pressure satisfied his equations. He applied the theory of probability to liquids near their freezing point, and found that in a liquid state there are some particles corresponding to the solid state and in a solid state some particles corresponding to a liquid state. What distinguishes a solid from a liquid is not, according to this theory, the fact that all particles are either solid or liquid, but simply that the rate of solidification or of liquefaction is in excess of the other process. The existence of an indefinite number of these small solid particles would have a marked influence on volume changes provided they were different in volume from those of the liquid state. But water expands on solidification, so that if the number of solid particles increases with decrease in temperature, the liquid would expand on cooling.

Whiting continues,   .

It may be allowable to suggest, that from almost any point of view, there will be in melting ice, only from 50 to 70 per cent of solid particles, and in freezing water nealy one-half as many, but in boiling water not more than one-third; so the number which disappear in melting is not more than twice the number which are eliminated when the liquid is raised to boiling. The expansion, therefore, from 0° to 100° instead of being 1.04, is probably from 0.08 to 1.10 and the real coefficient of expansion at 0° is from 0.0006 to 0.0008, increasing regularly with the temperature, as in the case of any ordinary liquid.

In those early days Julius Thomsen (3) concluded that water molecules are twice as heavy as vapor molecules, as did Raoult (4) from a study of freezing points of solutions. Armstrong (5) in 1888 advanced the idea that liquids in general are probably made up of complexes of the fundamental gas molecules through the action of residual affinity.   .

Vernon (6), quite independently, explained the temperature of maximum density by the presence of "water molecules aggregating together," and possessing a density smaller than water but larger than ice molecules. The water molecules were supposed to be $(H_2O)_2$ while the complexes were given a formula

$(H_2O)_4$. He also demonstrated from specific heat data that the increase in complexity of the molecules is accompanied by an evolution of heat.

To Röntgen (7) is commonly ascribed the first suggestion of "ice molecules" in liquid water. His paper followed that of Whiting by eight years but was very much more extensive in explaining the properties of this interesting liquid. He postulated that water is a saturated solution of ice molecules and that the concentration depends on the temperature, a decrease in temperature favoring the formation of the more complex ice molecules. He assumed further that the change in molecular state from complex to simple molecules, corresponding to melting of ice, has the result of decreasing the volume. On these assumptions he explained the point of maximum density. If water below 4° is heated, the volume would be changed by two actions; first there would be contraction due to the breaking up of the bulky complex molecules, and secondly there would be an expansion due to thermal expansion of the liquid molecules. If the former is in excess of the latter, the liquid would contract on heating; but if the latter is in excess of the former, there would be an expansion on heating. This is what is assumed to occur below and above the point of maximum density.

Röntgen said, from the analogy of ordinary saturated solutions, that pressure would decrease the number of ice molecules, and so cause a contraction. For a given pressure, this would be greater the lower the temperature. The compressibility of a normal liquid is smaller the lower the temperature. So it follows that for water there is a point of minimum compressibility at some temperature above the freezing point. Such a point is found at about 50°.

A third anomaly was the fact that the thermal coefficient of expansion of water at pressures of 3000 atmospheres is opposite in sign to that of other liquids, for it increases with increasing pressure. Normally this would decrease with increasing pressure, as does the compressibility. If the pressure breaks up the complex molecules fast enough for the resulting increase of volume to exceed the decrease due to the normal action of pressure, then

the thermal coefficient would increase with increasing pressure, as it does in the case of water.

Röntgen claimed that on the assumption that the number of ice molecules is decreased by pressure, the maximum density of water would occur at a lower temperature under pressure (as found by Amagat and others) and the freezing point would be lowered by pressure. Furthermore, within a certain interval of temperature, water subjected to an increase in pressure, would cause a cooling, on the assumption that the change of ice molecules to simple molecules uses up heat.

The last anomaly explained by Röntgen was the decreasing viscosity of water with increasing pressure. This was accomplished by the assumption that the simple molecules had a smaller viscosity than the ice molecules.

These postulates have been the foundation of practically all of the theories concerning the molecular state of water.

THE CONCEPT OF EQUILIBRIUM

The principles of thermodynamics were applied to the problem by Van Laar (8) who, after having accepted Ramsay and Shields' values for the quantitative degree of association, concluded that an equilibrium exists between double and single molecules, their proportions changing with conditions as postulated by Röntgen. Furthermore, when a second material like alcohol is added, he supposed that some of the double molecules are broken into single molecules, with a resultant contraction.

Röntgen's postulates were used by de Coppet (9) and Witt (10) to explain the displacement of the temperature of maximum density in solution. Witt also attempted to explain heats of solution, abnormal lowerings of the vapor pressure and excessive increases of osmotic pressure as they changed with concentration. He believed that this simple water molecule is $(H_2O)_2$ while the ice molecule is $(H_2O)_8$.

Sutherland (11), in 1901, presented a very interesting attempt to determine the relative amounts of the complex molecules under various conditions. His calculations led him to believe that water is a binary mixture of trihydrol $(H_2O)_3$ and dihydrol

$(H_2O)_2$ existing in a dynamic equilibrium.   Due to the hexagonal structure of ice, the solid molecules (those having the greater volume) were supposed to be trihydrol, $(H_2O)_3$.   From the density of ice and the usual expansion resulting upon melting the solid of a normal liquid, he calculated the probable density of the trihydrol molecules in the liquid state.   The change in density of the simpler molecules with a change of temperature was assumed to be given by the tangent drawn to the density-temperature curve at 100°.   Then by the use of Mendeléeff's equation for the expansion of normal liquids and the mixture law he calculated the various amounts of the different molecular species existing at any temperature.   At 0° the percent of $(H_2O)_3$ was 37.5, at 20° 32.1 per cent, and at 100° 21.7 per cent.   At the critical temperature the liquid was supposed to be nearly pure dihydrol. From this table of composition the various physical properties of the two components were calculated.   The change in composition with a change in pressure was also calculated, and resulted in the prediction of the pressure above which water would behave like a normal liquid composed of only one sort of molecule, dihydrol.

Sutherland reasoned further that if pressure causes dissociation of complex molecules, then surface tension would produce a layer of more highly associated molecules.   At temperatures below 40° this layer was supposed to be pure trihydrol.

The abnormally large heat changes accompanying a change in state are further evidence that the molecular nature of water is complex.   The heat of fusion must include the heat of reaction of di- to trihydrol, while the heat of vaporization must include the corresponding change from dihydrol to monohydrol (water vapor molecules).   Sutherland estimated each of these heats of reaction.

The decrease in viscosity of water with an increase of pressure or temperature was explained by the assumption that the smaller particles formed by the breaking up of the complexes have a smaller viscosity than the aggregates.   The fact that the viscosity of some dilute solutions is smaller than that of pure water was explained by the breaking up of some of the trihydrol under the action of the solute.

The fact that the freezing point occurs at a definite temperature even though the liquid consisted of di- and trihydrol in dynamic equilibrium gave Sutherland some difficulty. He suggested that this is caused by some sort of molecular resonance.

The changes in volume due to changes in polymerization have been considered by Richards (12) in connection with his lucid explanation of cohesive forces as they affect volume. Here the union of molecules was supposed to take place through the forces between oxygen atoms.

Hudson (13) explained the constant temperature of freezing by assuming that the ice molecules have a definite solubility in the water as well as a definite equilibrium concentration, both of these concentrations varying with the temperature. When they are equal (at 0° in pure water) solid appears. Above 0° the equilibrium concentration would be less than the solubility. If a second chemical substance is introduced, this is supposed to decrease the amount of ice molecules, making it necessary to cool the solution to a lower temperature before the equilibrium concentration reaches the saturation point of the ice molecules.

In 1908 Armstrong (14) attacked the problem from the chemical point of view and suggested the possibility of the existence of isomeric water molecules of the same molecular complexity but with different structure. These molecules have different activities depending upon their chemical structure in a manner similar to organic compounds. For example, the active isomers might be hydrol (OH_2) or hydronol $\left(H_2O \begin{smallmatrix} \diagup H \\ \diagdown OH \end{smallmatrix} \right)$ while the inactive molecules would be represented by closed systems with oxygen atoms joined to oxygens, each having a valence of four. These were called hydrones. He believed that in solutions of electrolytes the solute changed the proportions of these constituents and that the ions combined with the active forms of water. From this point of view he studied the volume changes on neutralizing acids and bases, the effects of salts on optical activity of sugar solutions, etc. More recent views will be discussed later.

In 1910 the Faraday Society (15) held a symposium on the structure of water. Papers were presented by Walden, Guye, Bousfield and Lowry, Sutherland, and Nernst. Walden showed that water is not an electrolyte except in solutions where its amphoteric nature allows it to form some kind of a "salt" with the other constituent. Guye presented his quantitative method of determining association in the liquid state, based on the assumptions that association existing in the vapor state proves association in the liquid, and that liquid water is made up of two sorts of molecules, a single and a double one. At 100° the association factor was equal to 1.86. Sutherland reiterated his belief that liquid water is a binary mixture of molecules, suggesting that any $(H_2O)_1$ was completely ionized into H^+ and OH^-.

Bousfield and Lowry presented an important paper in which they suggested that water is a ternary mixture, composed at low temperatures of $(H_2O)_3$ and $(H_2O)_2$, but at higher temperatures of $(H_2O)_2$ and $(H_2O)_1$. This conclusion was arrived at by considering water as the limiting case of a series of solutions of sodium hydroxide, where water is the most complex system, the complexity decreasing with increased concentration, until at 12 per cent the curves which represent changes in solution volume with temperature could be represented by a parabolic formula and at 42.5 per cent by a linear function over the temperature range 0° to 100°. The straight line, which is the limit of these curves, is tangent to the specific volume temperature curve for water at about 30°. (The solution volume of a solute is defined as the increase in the volume of the liquid which takes place when 1 gram of the solute is dissolved in 100 cc. of the solvent.) At 60° the specific volume curve for water begins to become abnormal, which was explained by the presence of increasing amounts of steam molecules. So Bousfield and Lowry differed from Sutherland in thinking that the line showing the normal behavior of pure water molecules is tangent to the actual curve at 30° rather than at 100°.

The solutions studied were those of sugar, acetic acid, and such

strong electrolytes as sodium hydroxide, silver nitrate, lithium chloride, etc. In such solutions the possibility of hydrate formation must be considered. In the curves obtained for the salts, there are decided maxima, which flatten out as the concentration of solute decreases. In such cases the curve for water is the simplest of the family. Hydration of the solute was used to explain these curves.

The last paper presented at the symposium was by Nernst, on the specific heat of steam, water, and ice. He showed that the data could be explained by accepting the expression:

$$2H_2O = (H_2O)_2 + 2519 \text{ calories}$$
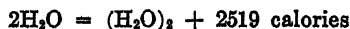
At the end of the discussion, Professor James Walker, the chairman of the meeting said:

I should think as a result of this discussion, one will soon find even in the textbooks that while ice is trihydrol, and steam monohydrol, liquid water is mostly dihydrol with some trihydrol in it near the freezing point and a little monohydrol near the boiling point.

Bousfield has used his theory of the structure of liquid water in several later papers. In the first of these (16) he pointed out that there is probably an intimate connection between the vapor pressure of water and the proportion of steam molecules in the liquid; that there was little doubt that the proportion of steam as well as ice molecules is reduced by the solution in water of any solute, and that this is connected with the reduction of the vapor pressure. He again pointed out (17) the "remarkable fact that steam molecules, like ice molecules, must be considered as bulky molecules." In the second paper (18) osmotic pressure was attributed to the thermal agitation of the vapor molecules $(H_2O)_1$. The addition of a solute was said to be accompanied by a shift in the equilibrium conditions of the liquid water, which results in the depression of the vapor pressure and of the freezing point.

CRYOSCOPIC DETERMINATIONS

Various other methods of attack have been used in attempting to determine the molecular state of water. One that would come

to mind immediately would be that of cryoscopic determinations (19). On the basis of the theories outlined, the assumption that the molecular state in pure liquid would be the same as that existing in the presence of a large excess of a second substance is false. The apparent molecular weight of the water, of course, varies with the concentration of the solution. With solvents such as p-toluidine, phenol, bromoform, methyl oxalate, ethylene bromide, and veratrol the calculated molecular weight of the water varies from 17.6 to 35.9. Bruni and Amadori (19) concluded that in nondissociating solvents water forms complex molecules and if double molecules exist exclusively, they do so only in very concentrated solutions, while in all solvents water tends to form simple molecules of monohydrol in dilute solutions. Oddo and Scandola (20) in 1910 critically reviewed the cryoscopic determinations and decided that in practically all solvents the water exists as $(H_2O)_2$ unless it combined with the solvent.

From a study of the system: water, ether, and succinic acid, Forbes and Coolidge (21) estimated that the association factor of the water dissolved in the ether is a little less than two.

THE CRYSTALLINE HYDRATES

Still another method of attack has been the study of crystalline hydrates, there being a supposed relation between water of crystallization and liquid water. One of the most common of the early attempts was to calculate the density of the water held in the crystal and then classify it as a particular form of hydrol. Pickering (22) said that the density is the same as that of ice, believing that ice is not an aggregate of the water molecule but an entirely different compound. A study of specific heats of hydrates confirmed his views. Thorpe and Watts (23) said that water of crystallization has a density of 1.24. Sutherland (24) calculated that the density of water held in $Li_2SO_4 \cdot H_2O$ is 1.31 and thus about the average of 31 other hydrates, as determined by Clarke (25). As a consequence, he assumed this to be the density of solid monohydrol. Biltz (26) has recently calculated the molecular volume of the water in $CuSO_4 \cdot 2H_2O$ to be 13.7 cc. Rosenstiehl (27) made an extensive study of some 179 hydrated

salts and concluded from the numbers of water molecules lost in each step of hydration that liquid water is a ternary mixture of $(H_2O)_1$, $(H_2O)_2$, and $(H_2O)_3$.

The x-ray analysis of hydrates shows the futility of these attempts. What data there are would show (28) that the water is present as H_2O units. In the hexahydrate of zinc bromate, for instance, the six H_2O groups are all equivalent and have a similar arrangement, probably about the zinc atom. In the alums the twelve water molecules fall into two groups of six equivalent molecules, probably arranged about the metal atoms.

THE STRUCTURE OF ICE

We have seen in the preceding discussion that the molecular unit of ice has been assumed to be trihydrol. This is affirmed by Fielding (29) while Duclaux (30) has suggested that its composition is between $(H_2O)_6$ and $(H_2O)_{23}$.

The researches of Tammann (31) and Bridgman (32) on water under high pressure and the evidence for believing the existence of several forms of ice are well known. They have studied water under high pressures over a large range of temperature and have found that there are five different kinds of ice, each one stable under particular conditions of pressure and temperature. Tammann believes that it exists in several forms which fall into two groups: (a) those which are lighter than liquid water, and (b) those which are heavier than liquid water. He concludes that ice belonging to group (a) separates only from liquid water rich in polymolecules, while those belonging to class (b) form from water containing an abundance of simple molecules. Ordinary ice, called I, belongs to group (a). Ice II and III he concludes to have the same form and to fall in class (b). He believes that liquid water under pressure behaves like a two component liquid.

In the papers on water, mentioned above, Bridgman discussed many of the abnormalities of water which change in magnitude with pressure and temperature. It was shown that water passes from an abnormal to a normal liquid as the pressure and the temperature increase; for instance, the minimum of compressibility at 50° is eliminated as the pressure increases. The point

of maximum density, which is at 4° at atmospheric pressure, is depressed rapidly by increasing pressure until it has fallen below the freezing point at 300 kgm. per square centimeter. These abnormalities of changes in volume with pressure and temperature were explained on the basis of polymerization, the presence of only two kinds of particles being assumed.

At temperatures near the melting point Tammann (33) believes that diffusion in a crystal occurs so that the molecular weight of a crystal near its melting point may be discussed. He calculated from Walden's (34) rule, admitting the doubtfulness of its application, that the molecular formulas of the ice I, III, IV, and VI are $(H_2O)_3$, adding that the differences in the ices are due to isomerism rather than polymerism. He further concluded that only such molecules form crystals as exist in the liquid.

With the application of x-rays to the study of crystals of polar substances our concept of molecular weight in the solid state has little significance, giving place to the well known lattice structure in which molecules are indiscernible. The x-ray analyses of ice have yielded divergent results. A Laue photograph by Rinne (35) assigned ice to the hexagonal system with an axial ratio of $a : c = 1 : 1.678$. From an assumption that the crystals are not twinned, Gross (36) found that a unit cell containing two molecules of H_2O with axial ratio of 1.60 is compatible with the Laue photograph. According to the spectrometric results of St. John (37), the unit cell contains four molecules of H_2O with a ratio $a : c = 1 : 1.4026$. The powder photographs of Dennison (38) have been interpreted to yield a unit of yet another size, having an axial ratio of 1.62. The close approximation of this axial ratio to the ratio for the closely packed grouping of spheres (1:1.633) was taken as an indication that the molecules of water are associated into $(H_2O)_2$ groups (39) which are themselves closely packed.

These data of Dennison have been used to confirm the structure of Bragg (40), arrived at by independent calculation. Bragg believes that no molecular unit exists in ice, that each oxygen atom is at the center of gravity of four neighboring oxygen atoms, from each of which it is separated by a hydrogen atom. In

organic liquids (41) two, three, or four molecules may exist as such in a crystal.

Wyckoff (42) in reviewing these determinations concluded that there is such serious conflict among the reported experimental data that "consequently nothing definite can be considered as known about its atomic arrangement."

VAPOR DENSITY EVIDENCE

Until recently it has been believed (43) that water in the vapor state is composed principally of single molecules but that there is a certain quantity of double molecules also. In fact, Guye's (44) method of determining the extent of association of a liquid was dependent upon the concentration of associated molecules in the vapor. Bose (45) calculated from the density determinations of Kornatz the concentration of double molecules to be about 10 per cent from 0° to 200°, while Oddo (46) assumed dissociation below 32° C. and association above that temperature in order to explain the data (41.4 per cent of $(H_2O)_2$ at 270°).

Recently, however, both Kendall (47) and Menzies (48) have concluded from a recalculation of the old vapor density values that there is no evidence for the existence of double molecules when the densities are corrected for deviations from the perfect gas law. Shirai (49) confirmed this conclusion for the temperature range 80° to 140°. However, new determinations of density with an accuracy within 0.1 per cent at atmospheric pressure and at a temperature range from 98° to 200° have been reported by Maass and Mennie (50). The results show greater divergence from the ideal density than can be accounted for on the basis of the equation of state. As a consequence they have adopted the hypothesis of polymerization with the formation of double molecules. This association in the vicinity of 100° and 1 atmosphere pressure was estimated to be less than 0.9 per cent.

Gillet (51) has attempted to extend the fundamental portions of the theory of the polymerization of water to both real and colloidal solutions, but made many unreasonable assumptions.

RECENT THEORIES

H. E. Armstrong has long been a proponent of the marvelous nature of water and has pointed out on every possible occasion the neglect which chemists have shown in considering it a mere inert solvent. The development of his views concerning water and his attack on the theory of ionization since 1888 are to be found in his very recent collection of essays, "The Art and Principles of Chemistry" (52). He believes that water is a complex mixture saturated with the "gas" hydrone, OH_2, which may become active under the influence of a dissolved substance. The relative proportion of this molecular species, either in water or an aqueous solution, is supposed to be measured by the vapor pressure. In a solution the hydrone molecules will be "distributed" upon the solute, forming $M\begin{smallmatrix}\diagup H\\\diagdown OH\end{smallmatrix}$ if the solute is a non-electrolyte (M), or $RX\begin{smallmatrix}\diagup H\\\diagdown OH\end{smallmatrix}$ and $H_2O\begin{smallmatrix}\diagup R\\\diagdown X\end{smallmatrix}$ when the solute is an electrolyte (R^+X^-). As the concentration of the electrolyte decreases, $RX\begin{smallmatrix}\diagup H\\\diagdown OH\end{smallmatrix}$ is supposed to be converted into hydronol, $H_2O\begin{smallmatrix}\diagup H\\\diagdown OH\end{smallmatrix}$, until ultimately the solution contains the solute only in the form $H_2O\begin{smallmatrix}\diagup R\\\diagdown X\end{smallmatrix}$ together with an equal number of molecules of hydronol. These "distributed" complexes are the active constituents of the system and have the power to attract single hydrones, thus serving to restore the hydrone equilibrium. The osmotic pressure is supposed to be due to these extra hydrone molecules, the latter being proportional to the number of "distributed complexes."

A theory similar to that of Armstrong has been presented by

Kling and Lassieur (53). The hydrol (Armstrong's "hydrone," and the simple H_2O molecule) is supposed to be a conductor of electricity while the polymers are not. This hydrol is believed to exist in two tautomeric forms, $H_2 = O$ and $H-O-H$, the first basic and the second acidic, $(H_2O)_2$ being the neutralized product. When an acid is added, it is supposed to combine with $H_2 = O$. The ionic conception of hydrogen ion concentration is substituted by the concentration of the acid hydrol molecules.

Both of these theories are considered inadequate by Auger (54). If there is an equilibrium between $(H_2O)_n$ and nH_2O, any hydrone that was fixed by a solute would be replaced by the equilibrium and as a consequence, the vapor pressure (supposed to be directly proportional to the concentration of hydrone) would not be lowered. His objection to the second theory is that there can be no equilibrium between the two tautomeric forms of hydrol or otherwise there would never be an excess of one form in solution.

Tammann (55) has recently contributed a paper on the molecular composition of water, reiterating that the fact that water reaches a minimum volume at 4° can be explained by the presence of molecules of greater volume, which increase in concentration as the temperature is decreased. This molecular form (Type 1) is supposed to have the same space lattice as ordinary ice and to exist at temperatures up to 50° and pressures between 0 and 2500 kgm. per square centimeter. Other forms of water molecules are supposed to be present in the liquid, but they are not so important in determining the physical properties. The degree of polymerization of this form has been determined from thermodynamic data to be either $(H_2O)_9$, in which case it splits into $9(H_2O)_1$, or $(H_2O)_6$, in which case it divides into $2(H_2O)_3$. The heats of dissociation and specific heats are estimated. The change in viscosity with temperature and pressure does not seem to be strictly dependent upon the concentration, but surface tension and index of refraction below 60° are proportional to the concentration of Type 1 molecules.

In a second paper Tammann (56) pointed out that the addition of a relatively nonvolatile substance (salts) displaces the volume minimum to lower temperatures, and decreases the compressi-

bility in the same manner as an increase of external pressure. These alterations were attributed to changes in molecular complexity of the water, to a decrease in the concentration of molecular Type 1. If a non electrolyte is present whose solubility increases with the amount of Type 1 (more soluble in cold water than in hot), then it will be less soluble in the salt solution. This idea was tested by solubility determinations.

X-RAY ANALYSIS OF LIQUID WATER

An entirely new method of studying liquids has come from the application of x-rays to the liquid state. When a liquid is substituted for the solid in a powder photograph experiment a very different diffraction effect is produced, the picture consisting of broad but distinct bands rather than sharp lines. The bands are similar to those produced by glassy solids. There is at present no agreement as to the source of these bands.

Debye and Scherrer (57) proposed that the bands arise from interference of rays scattered by the atoms within the chemical molecules of the liquid. Hewlett (58) assumed a crystal structure in the liquid state. On the other hand, Keesom and de Smedt (59) believed that the bands arise from rays scattered by molecules that are arranged in a more or less regular manner. On the assumption that the molecules are closely packed spheres, de Smedt (60) has calculated the degree of association of several organic liquids. Raman (61) does not agree with these conceptions but believes that the "liquid patterns" arise from regular differences in density existing in the liquid. An attempt has been made by Raman and Ramanathan (62) to calculate this regularity thermodynamically from the compressibility of the liquid.

A series of experiments by Wyckoff (63) on liquid mixtures shows that the pattern of a liquid mixture is the sum of the diffraction of its components and substantiates the conclusion that the origin of the pattern is within, rather than between molecules, although the results do not exclude the possibility of their arising from characteristic association of the molecules. Katz (64) has shown that the degree of polymerization of a solute does not affect the diameter of the diffraction ring, for such sub-

stances as rubber in isoprene, etc. Zernike and Prins (65) demonstrated to their satisfaction that the patterns cannot be due to arrangements of electrons in atoms or of atoms in molecules. Future work alone will give the true explanation of this property of liquids.

Langmuir (66) believes that a liquid resembles a solid in that there is no molecular unit of structure, but rather that the molecules are held together by chemical forces of the same character as the forces acting between atoms. Recently Daniels and Williams (67) and Antonoff (68) have published data on the specific heats of liquids which seem to show that this property is discontinuous. This is explained by probable changes in molecular complexity.

Latimer and Rodebush (69) have studied the structure of water from the electronic point of view and consider it to occupy an intermediate position between hydrogen chloride and ammonia. They suggest that a free pair of electrons on one molecule might be able to exert sufficient force on a hydrogen atom held by a pair of electrons on another molecule to bind the two molecules together. Such a union is not limited to the formation of double or triple molecules. This sort of association is very different from that of acetic acid (in which definite double molecules are supposed to be formed) and is probably the factor that produces the extremely high dielectric constant.

AGGREGATES IN EQUILIBRIUM

The hypothesis that liquid water contains various aggregates in equilibrium and that the equilibrium is changed by a second substance has found a number of applications within the past decade. In a very interesting paper by Richards and Palitzsch (70) it is shown that the solution volumes, viscosities, surface tensions and compressibilities of aqueous solutions of urethane can all be explained by assuming that the bulky trihydrol molecules (according to the theory of Bousfield and Lowry) are broken up as an effect of the solute. The compressibility curve, for example, shows a very decided minimum, the compressibility of the dilute solutions (up to about 25 per cent) being less than that

of pure water. The decrease is attributed to the effect of the solute on the solvent, the bulky polyhydrol molecules breaking up to form molecules having a smaller volume and compressibility, while the increase in the property at higher concentrations is attributed to the urethane possessing a greater compressibility than the dihydrol[1] which has increased in amount in the liquid as the concentration of urethane has increased.

Pagliani (71) has used the same explanation for solutions of alcohol and water, although the viscosities could not be explained as easily as the compressibilities.

The viscosities of certain dilute aqueous solutions are less than that of pure water at room temperatures. To this has been applied the misnomer of "negative viscosity." A few alcohol and glycerine solutions show the same behavior. Rabinovich (72) has made a critical study of the different factors acting upon the viscosity, particularly those factors which are able to lower the internal friction of the solvent. One of the most important of these factors is supposed to be the depolymerization of the associated solvent and it is this factor alone which is able to produce by itself negative viscosity. The simpler water molecules are supposed to have a viscosity very much smaller than the polyhydrol, so that the change in viscosity through a change in solvent more than offsets the increase due to the presence of a viscous solute.

POLYMERIZATION

Richards and Chadwell (73) contributed further evidence for the theory of the polymerization of water through a study of volume changes and compressibilities of aqueous solutions of non electrolytes. The results were explicable by reference to three causes: (a) the mutual affinity or attraction manifested by the liquids for one another in relation to the cohesive affinities of the pure liquids; (b) the effect of depolymerization of one or both liquids, and possible solvation; (c) the effect of the several compressibilities of the cohering substances.

[1] A possible structure for dihydrol is given by Anderegg, Proc. Indiana Acad. Sci. (1923), p. 93.

An illustration of the application of these considerations is of interest in our present discussion. It was found that the contractions taking place upon the formation of one liter of aqueous solutions of urethane, methyl acetate, and ethyl ether increased with an increase in concentration of the solute and for a given molal concentration was greater for ether than for methyl acetate, which in turn was greater than urethane. These three solutes are considered to be little associated.

They are, therefore, suitable for preliminary comparison. The average compressibilities (between 100 and 300 megabars) of these three substances in the liquid condition at 20° are, respectively, 132, 88 and about 46 (each $\times 10^{-6}$). Evidently, the contractions (15, 8 and 4.2 cc.) which take place on forming a solution containing one mole of solute per liter are roughly proportional to these compressibilities.

One might infer that the compressibility of the solute is the only factor in the volume change, but this inference would be superficial. It is not the compressibility of the solute alone which must be considered, but rather its relation to that of the solvent. Now the compressibility of liquid urethane is not far from that of water, although probably somewhat greater. If no other circumstance entered into the situation, liquid urethane ought to be nearly "isofluid" with water, involving no volume change on mixing. There is thus reason to believe that the rather large volume change which actually occurs when urethane is dissolved in water is primarily due not to further compression of urethan or water in the act of solution, but rather to some other circumstance, presumably the depolymerization of some of the water, which would cause a diminution in volume, since there can be little question that the more complex molecule of water is more bulky than a less complex molecule. This conclusion gives a clue which will be followed later as to the extent of polymerization of water. It does not, however, invalidate the conclusion that compressibility as indicated by the behavior of the solutions of ether, methyl acetate and urethane, is probably an essential factor in the volume change, the later being greater, the greater the compressibility of the solute. The same solvent is common to all.

The relative differences in contraction cannot be due to the effect of several affinities because, judging from the extent of solubility, ether has the least and urethane the greatest affinity of the three substances for water. This latter inference might also be drawn from the heats of

solution of similar substances in various solvents, determined by Speyers.

Turning now to a study of the effect of a single solute on different solvents, the changes in volume were determined for the solution of urethane in benzene, alcohol, water, and ether.

Again ether, the most compressible of all these solvents, gives by far the greatest change in volume. In the case of benzene the smaller compressibility (77×10^{-6}) and small affinity (shown by the slight solubility and great negative heat of solution) are presumably the reasons why this substance gives a slight *increase* rather than a decrease in volume. Alcohol and water behave as would be expected, taking account of their association; the volume change in the case of water is over twice as great as in the case of alcohol, although the compressibilities show the opposite relation. When urethane is dissolved in water, the dissociation of a part of the more complex molecules of water may be assumed to cause considerable contraction, and this contraction is to be added to that (if any) due to the mutual compression of the two substances. On the other hand, in the case of alcohol the dissociation (by causing expansion) would tend to decrease the volume change. Hence, the transposition of the two curves is only to be expected. The effect of change of polymerization may then be inferred (with regard to this particular pair) to exceed that due to the different compressibilities.

These qualitative considerations are inevitably incomplete, especially in view of the fact that the compressibilities of all substances diminish (to various extents) with increasing pressure. Nevertheless, they are not without significance.

No evidence as to the actual number of water molecules in the polyhydrol was found in these experiments, although an approximate estimate was made of 28 per cent of polymer present in water at 20°.

One of the most impressive pieces of evidence for the theory of the polymerization of water was found in the fact that the compressibilities of aqueous solutions of ether and methyl acetate are less than water, even though the compressibilities of the pure solute are very much greater. (The coefficient of compressibility of ether is about three times and that of methyl acetate two times

as great as that of water.)   The half molal solutions of these
solutes and urethane possess a compressibility about 3 per cent
less than that of pure water.   It appears that the compressibility
of any dilute aqueous solution is less than that of water, and that
the only plausible cause of this common effect is the depolymeri-
zation of water.

Since the viscosity of the depolymerized water molecules is
supposed to be less than that of the polyhydrol, the viscosities
of these aqueous solutions might be expected to be considerably
lower than that of water.   Chadwell has shown that this is not
the case at a temperature of 25°, (74), but instead the viscosities
of the solutions are greater.   This increase in viscosity seems to
be a general property of aqueous solutions of nonelectrolytes, at
least at a temperature of 25°.   The effect of the change in poly-
merization is covered up by other factors, possibly change in
volume, etc.

Still another abnormal property that can be explained by the
theory is the variation with temperature of the magnetization
(75).

An entirely different application of the theory is the explana-
tion of Bancroft (76) for the peptization of gelatin by various
salt solutions.   The addition of salt affecting the water equilib-
rium is supposed to affect the peptization if one of the forms of
water is the determinant favoring peptization.   One of Bancroft's
students, Bowe (77), has used the theory in the study of the
neutral salt effect.

The pertinence of the water equilibrium to phenomena related
to aqueous solutions of electrolytes has been recognized by many
investigators, but very little progress has been made in applying
the ideas to the electrical properties of solutions.   For instance,
Kendall (78) has considered the consequences of a shift in equilib-
rium in an elucidation of the application of ideal solution
equations to dilute aqueous solutions.

Bancroft (80) in reviewing the present status of the theory of
electrolytic dissociation says:

Forty years of intensive development have brought us to the point
where we cannot determine any electrolytic dissociation with any

degree of accuracy and where we question the significance of the term "electrolytic dissociation." . . . . It is easy enough to point to one factor which has been neglected practically completely and which may be the one which has caused most—and perhaps all—of our difficulties. For years H. E. Armstrong in England has chided the physical chemists for considering water only as water, whereas it is a complex and variable mixture. This criticism seems well founded; but unfortunately, Armstrong has never succeeded in showing what could be done with his idea and consequently, the idea has been valueless hitherto. Everybody admits that water is a polymerized liquid and that the degree of polymerization may change on the addition of electrolytes. Sutherland, Lewis, McBain and others have suggested such a displacement of equilibrium as a possible source of error in our physical chemistry calculations; but nobody seems to have made a serious attempt to see how adequate this suggestion is.

Let us hope that in the near future rapid progress will be made in this important field of investigation.

BIBLIOGRAPHY

(1) TURNER: Molecular Association, Monographs on Inorganic and Physical Chemistry edited by Findlay. Longmans, Green & Co., (London, 1915). This monograph reviews the subject up to 1915. The Ramsay and Shields method of determining association has been criticized by Bennett: Jour. Chem. Soc., 125, 958 (1924); Maass in Treatise on Physical Chemistry, edited by H. S. Taylor, Van Nostrand Co., New York City (1925), Vol. I, p. 134. For more recent work on the quantitative determination of association the following references may be cited: CARDOSO AND BATTISTA: Anales soc. espan. fis. quim., 20, 420 (1922); DRUCKER: Z. physik. Chem., 68, 616 (1919); FIELDING: Chem. News, 122, 13, 289 (1921); 123, 1 (1921); GAPON: Jour. Chim. Ukraine, 1, 481 (1925); Chem. Abs., 20, 3104 (1926); GARNER: Science Progress, 19, 382 (1925), A Review; GARVER: J. Phys. Chem., 16, 454, 471, 669, 679 (1912); 17, 386 (1913); 19, 500 (1915); KARPEN: Bull. sec. sci. acad. Roumaine, 8, 22 (1924); Chem. Abs., 19, 758 (1925); MACLEOD: Trans. Faraday Soc., preprints, 1923, 1925; Trans. Faraday Soc., 20, 348 (1924); ROCARD: J. phys. radium, 6, 198 (1925); SCHULZE: Physik. Z., 26, 153 (1925); SCHUSTER: Ber., 58B, 2183 (1925); Z. anorg. allgem. Chem., 146, 299 (1925); SWIENTOSLAWSKI: Roczniki Chem., 1, 276 (1921); through Chem. Abs., 16, 3025 (1922); TYRER: J. Phys. Chem., 19, 81 (1915); WALKER, E. E.: Phil. Mag., 47, 111, 513 (1925).

(2) WHITING: A Theory of Cohesion, Harvard University, Cambridge, Mass., (1884).

(3) THOMSEN: Thermochemischen Untersuchungen, Vol. III (1883), page 181; Ber., 18, 1088 (1885).

(4) RAOULT: Ann. chim. phys. (6), 4, 401 (1885).

(5) ARMSTRONG: J. Chem. Soc., 53, 131 (1888).

(6) VERNON, H. M.: Phil. Mag., 31, 387 (1891).

(7) RÖNTGEN: Ann. phys. chim. (Wied.), 45, 91 (1892).

(8) VAN LAAR: Z. physik. Chem., 31, 1 (1899).

(9) de COPPET: Compt. rend., 115, 606 et seq. (1893); Wright, J. Chem. Soc., 115, 119 (1919).

(10) WITT: Öfver K. Vetens. Akad. Forhandl., 57, 63 (1900). Through J. Chem. Soc., 80, ii, 498 (1901); and Chem. Zentr., 71, 946 (1900).

(11) SUTHERLAND: Phil. Mag. (V) 50, 460 (1901).

(12) RICHARDS: Proc. Am. Acad. Arts Sci., 39, 594 (1904).

(13) HUDSON: Phys. Rev. (1), 21, 16 (1905).

(14) ARMSTRONG: Proc. Roy. Soc. London, 81A, 80 (1908).

(15) The Constitution of Water, a General Discussion. Trans. Faraday Soc., 6, 71 (1910).

(16) BOUSFIELD: Proc. Roy. Soc. London 90A, 41 (1914).

(17) BOUSFIELD AND LOWRY: Proc. Roy. Soc. London, 74, 280 (1904). Phil. Trans., 204A, 282 (1905).

(18) BOUSFIELD: Trans. Faraday Soc., 13, 141 (1917).

(19) SUTHERLAND: Phil. Mag. (6), 12, 1 (1906); EYKMAN: Z. physik. Chem., 4, 497 (1889); BRUNI AND AMADORI: Trans. Faraday Soc., 5, 290 (1910); Gazzetta, 40, i, 1 (1910); JONES AND MURRAY: Amer. Chem. J., 30, 193 (1903); ROZSO: Z. Elektrochem., 17, 934 (1911).

(20) ODDO AND SCANDOLA: Gazzetta, 40, ii, 172 (1910).

(21) FORBES AND COOLIDGE: J. Amer. Chem. Soc., 41, 150 (1919).

(22) PICKERING: J. Chem. Soc., 95, 127 (1909).

(23) THORPE AND WATTS: J. Chem. Soc., 37, 115 (1880).

(24) SUTHERLAND: Trans. Faraday Soc., 6, 104 (1910).

(25) CLARKE: Chem. Zentr., 6, 146 (1875).

(26) BILTZ: Z. anorg. allgem. Chem., 143, 231 (1925).

(27) ROSENSTIEHL: Compt. Rend., 152, 598 (1911); Bull. Soc. chim. (4), 9, 281, 284, 291 (1911).

(28) WYCKOFF; The Structure of Crystals, Chemical Catalog Co., New York City (1924), p. 396.

(29) FIELDING: Chem. News, 122, 13 (1921).

(30) DUCLAUX: Compt. Rend., 152, 1387 (1911).

(31) TAMMANN: Z. physik. Chem., 72, 609 (1910).

(32) BRIDGMAN: Proc. Am. Acad. Arts Sci., 47, 441 (1912).

(33) TAMMANN-MEHL: The States of Aggregation, Van Nostrand Co., New York City (1925), p. 147.

(34) WALDEN: Z. physik. Chem., 82, 172 (1913).

(35) RINNE: Ber. K. Sachs. Ges. Wiss. Mat.—Phys. Klasse, 69, 57 (1917).

(36) GROSS: Centr. Win. Geol., 1919, p. 201.

(37) ST. JOHN: Proc. Nat. Acad. Sci., 4, 193 (1918).

(38) DENNISON: Phys. Rev., 17, 20 (1921); Chem. News, 122, 54 (1921).

(39) LANGMUIR: Science, 53, 580 (1921). Through Chem. Abs., 16, 3778 (1922).

(40) BRAGG: Proc. Phys. Soc., London, **34**, 98 (1922).   Bragg's crystal model is used by Lowry and Burgess, J. Chem. Soc., **123**, 2111 (1923); Lowry, Trans. Faraday Soc., **20**, Part III (1924); Huggins, Phys. Rev., **21**, 205 (1923).

(41) BRAGG: J. Chem. Soc., **121**, 2766 (1922).

(42) WYCKOFF: The Structure of Crystals, Chemical Catalog Co., New York City (1924), p. 266.

(43) MELLOR: Treatise on Inorganic and Theoretical Chemistry, Longmans, Green & Co., London (1922), Vol. I, p. 460.

(44) GUYE: Trans. Faraday Soc., **6**, 78 (1910).

(45) BOSE: Z. Elektrochem., **14**, 269 (1908).

(46) ODDO: Gazzetta, **45**, i, 319, 395 (1915).

(47) KENDALL: J. Amer. Chem. Soc., **42**, 2477 (1920).

(48) MENZIES: Ibid., **43**, 851 (1921).

(49) SHIRAI: Bull. Chem. Soc. Japan, **2**, 37 (1927).   Through Chem. Abs., **21**, 1727 (1927).

(50) MAASS AND MENNIE: Proc. Roy. Soc. (London), **110A**, 198 (1926).

(51) GILLET, C.: Rev. gen. mat. color., **25** and **26**, six communications, 1921 and 1922.   Chem. Abs., **16**, 3019 (1922).

(52) ARMSTRONG: Essays on the Art and Principles of Chemistry, Jarrold & Sons, Norwich, England; Macmillan Co., New York (1927).   Also see Proc. Roy. Soc. (London), **103A**, 610 (1923).

(53) KLING AND LASSIEUR: Compt. rend., **177**, 109 (1923).

(54) AUGER: Compt. rend., **178**, 330 (1924).

(55) TAMMANN AND SCHEIL: Z. anorg. allgem. Chem., **158**, 1 (1926).

(56) TAMMANN AND BOTSCHWAR: Ibid., **158**, 25 (1926).   Also see TAMMANN: Über die Beziehungen zwischen den inneren Kräften und Eigenschaften der Lösungen, Leopold Voss, Hamburg and Leipzig (1907), p. 179.

(57) DEBYE AND SCHERRER: Nachr. Gesell. Wiss. Göttingen (1916), p. 16.   See WYCKOFF: The Structure of Crystals, Chemical Catalog Co., New York City (1924), p. 383.

(58) HEWLETT: Phys. Rev., **20**, 688 (1922).   See EASTMAN: J. Amer. Chem. Soc., **46**, 922 (1924).

(59) KEESOM AND DE SMEDT: Proc. Roy. Acad. Sci. Amsterdam. **25**, 118 (1922); **26**, 112 (1923); J. phys. radium, **4**, 144 (1923).

(60) DE SMEDT: Bull. classe sci. acad. roy. Belg., **10**, 366 (1924).

(61) RAMAN: Nature, **111**, 185 (1923).

(62) RAMAN AND RAMANATHAN: Proc. Indian Assoc. Cultivation of Science, **8** (2), 127 (1923).   Through Chem. Abs., **18**, 2839 (1924).

(63) WYCKOFF: Amer. J. Sci., **5**, 455 (1923).

(64) KATZ: Z. physik. Chem., **125**, 321 (1927).

(65) ZERNIKE AND PRINS: Z. Physik., **41**, 184 (1927).

(66) LANGMUIR: J. Amer. Chem. Soc., **38**, 2221 (1916).

(67) DANIELS AND WILLIAMS: J. Amer. Chem. Soc., **46**, 1570 (1924).

(68) ANTONOFF: Phil. Mag. (6), **50**, 265 (1925); (7) 1, 1121 (1926); (7) **3**, 571 (1927).

(69) LATIMER AND RODEBUSH: J. Amer. Chem. Soc., **42**, 1431 (1920).

(70) RICHARDS AND PALITZSCH: J. Amer. Chem. Soc., **41**, 59 (1919).   Reviewed by W. C. McLewis, Science Progress, **14**, 31 (1919).

(71) PAGLIANI: Gazzetta 50, i, 186 (1920).
(72) RABINOVICH: J. Amer. Chem. Soc., 44, 954 (1922). See also HERZ AND MARTIN: Z. anorg. allgem. Chem., 132, 41 (1924).
(73) RICHARDS AND CHADWELL: J. Amer. Chem. Soc., 47, 2283 (1925). For a discussion of aqueous solutions of electrolytes, see BAXTER: J. Amer. Chem. Soc., 33, 922 (1911); BAXTER AND WALLACE: Ibid, 38, 70, 259 (1916).
(74) CHADWELL: J. Amer. Chem. Soc., 48, 1912 (1926). See MacLEOD: Preprints Faraday Soc., 1923, 1925; Trans. Faraday Soc., 20, 348 (1924); SUGDEN: J. Chem. Soc., 125, 1177 (1924).
(75) COBRERA AND DUPERIER: J. phys. radium, 6, 121 (1925).
(76) BANCROFT: J. Phys. Chem., 30, 1194 (1926).
(77) BOWE: J. Phys. Chem., 31, 291 (1927).
(78) KENDALL: J. Amer. Chem. Soc., 43, 1391 (1921).
(79) BANCROFT: Review of Physical Chemistry in "A Half Century of Chemistry in America," Amer. Chem. Soc. (1926), p. 94.